La Natura della Comprensione Linguistica

Fenomenologia della Comprensione e Fondamenti Filosofici della Semantica

Il Paradosso Fenomenologico della Comprensione Immediata: Trasparenza Esperienziale versus Complessità Computazionale Sottostante

Quando un parlante competente della lingua italiana ascolta l'enunciato "il gatto è sul tappeto", l'esperienza fenomenologica che accompagna questo atto di comprensione possiede una qualità peculiare di immediatezza e trasparenza che sembra resistere a qualsiasi analisi ulteriore. Il significato appare semplicemente presente alla coscienza del parlante, senza mediazione apparente, senza sforzo introspettivamente rilevabile, senza traccia delle operazioni computazionali intermedie che devono necessariamente intervenire tra la percezione del segnale acustico o visivo e l'apprensione del contenuto proposizionale. La mente sembra cogliere direttamente il pensiero espresso dall'enunciato, come se vi fosse una sorta di "visione intellettuale" del significato che bypassasse completamente qualsiasi processo di decodifica o computazione simbolica. Non vi è alcuna consapevolezza fenomenologica delle cascate di processi neurali che operano al di sotto della soglia della coscienza, processi che si dispiegano a velocità dell'ordine dei millisecondi e che coinvolgono vasti network di popolazioni neuronali distribuite attraverso molteplici regioni corticali e sottocorticali del sistema nervoso centrale.

Questo carattere di immediatezza costituisce quello che potremmo definire il paradosso fondamentale e costitutivo della comprensione linguistica: la radicale incongruenza, l'apparente incommensurabilità, tra la semplicità fenomenologica dell'accesso al significato e la straordinaria complessità computazionale dei meccanismi neurali e cognitivi che rendono possibile tale accesso. William James, con la sua caratteristica acutezza nell'osservazione fenomenologica che lo distingueva come uno dei più grandi psicologi introspettivi del 19° secolo, aveva già identificato questo fenomeno nei suoi "Principles of Psychology" del 1890, introducendo la celebre espressione "the feeling of meaning" (la sensazione di significato) per designare quella peculiare qualità esperienziale di familiarità, riconoscimento e comprensione che accompagna l'elaborazione di espressioni linguistiche conosciute. Questa sensazione, James notava con grande perspicuità, è qualitativamente distinta dalla mera percezione di suoni articolati nel caso del linguaggio parlato, o dalla visione di forme grafiche sulla pagina stampata nel caso del linguaggio scritto: essa porta con sé un carattere intrinseco di intenzionalità, di essere diretta verso qualcosa che trascende le proprietà fisiche immediate del segnale linguistico stesso, qualcosa che Frege avrebbe in seguito teorizzato come il Gedanke (pensiero) espresso dall'enunciato.

La ricerca psicolinguistica contemporanea, armata di metodologie sperimentali di straordinaria precisione temporale e spaziale—tecniche che James non avrebbe potuto nemmeno immaginare—ha rivelato che questa immediatezza fenomenologica è, in larga misura, un'illusione cognitiva, o più precisamente, il prodotto finale emergente di una serie di elaborazioni neurali altamente complesse, gerarchicamente organizzate, e massicciamente parallele che procedono con tale rapidità ed efficienza automatica da risultare completamente opache alla coscienza introspettiva. Gli studi di eye-tracking durante la lettura, iniziati pionieristicamente da Émile Javal alla fine del 19° secolo ma portati a maturità tecnologica solo negli ultimi decenni, hanno documentato che i movimenti oculari non procedono in modo uniforme e continuo lungo il testo, come l'esperienza soggettiva potrebbe suggerire, ma consistono invece in una sequenza di movimenti saccadici rapidi, della durata tipica di 20-30 millisecondi, alternati a fissazioni durante le quali l'occhio rimane relativamente stazionario per circa 200-250 millisecondi. Già durante queste brevi fissazioni, il sistema cognitivo ha iniziato l'elaborazione lessicale e sintattica delle parole all'interno della finestra percettiva, estraendo informazioni ortografiche, fonologiche, morfologiche e semantiche in parallelo attraverso meccanismi neurali la cui architettura funzionale stiamo appena cominciando a comprendere nei suoi dettagli implementazionali.

Le tecniche di registrazione dei potenziali evento-correlati (Event-Related Potentials, ERP), che misurano l'attività elettrica cerebrale con risoluzione temporale dell'ordine del singolo millisecondo attraverso elettrodi posizionati sulla superficie dello scalpo, hanno permesso di dissezionare la dinamica temporale dell'elaborazione linguistica con un dettaglio cronometrico senza precedenti nella storia delle neuroscienze cognitive. Questi studi, iniziati negli anni '80 del 20° secolo con il lavoro pionieristico di Marta Kutas e Steven Hillyard, rivelano che il cervello inizia a distinguere parole da non-parole già entro 120-160 millisecondi dalla presentazione dello stimolo, come evidenziato dalla componente N1, una deflessione negativa del potenziale elettrico registrato sulla superficie dello scalpo. L'accesso al significato lessicale, riflesso dalla componente N400, si manifesta con notevole consistenza attorno ai 400 millisecondi dopo la presentazione della parola, mentre processi di integrazione sintattica più tardivi, di rianalisi di strutture ambigue e di riparazione di violazioni sintattiche si manifestano nella componente P600, che emerge tipicamente tra i 500 e gli 800 millisecondi. Tutto questo complesso macchinario computazionale opera al di sotto della soglia della consapevolezza cosciente, in quello che i neuropsicologi cognitivi chiamano il "processamento automatico preattentivo".

Il risultato finale che emerge alla coscienza è semplicemente l'esperienza unitaria di aver compreso, di aver colto il significato dell'enunciato in un atto apparentemente istantaneo e indivisibile. È come se il sistema cognitivo presentasse alla coscienza soltanto il prodotto finale della computazione, mantenendo nascosti tutti i passaggi intermedi, analogamente a come un computer moderno presenta all'utente solo l'interfaccia grafica finale senza esporre i miliardi di operazioni logiche a livello dei transistor e dei circuiti integrati che avvengono nei microprocessori. Questa architettura cognitiva ha un senso adattivo profondo quando la consideriamo dal punto di vista della teoria evoluzionistica della cognizione: se dovessimo essere consapevoli di ogni singolo passo dell'elaborazione linguistica—ogni accesso lessicale, ogni costruzione di struttura sintattica, ogni inferenza pragmatica—la comunicazione diventerebbe impossibilmente lenta e laboriosa, e le risorse attenzionali estremamente limitate della coscienza fenomenica verrebbero rapidamente sovraccaricate fino al collasso funzionale. L'illusione della immediatezza è quindi non un difetto ma una

caratteristica adattiva, funzionalmente necessaria per l'efficienza della comunicazione linguistica nel contesto delle pressioni selettive evolutive che hanno modellato l'architettura della cognizione umana nel corso di centinaia di migliaia di anni.

Tuttavia, questa illusione pone un problema epistemologico e metodologico fondamentale per la costruzione di una teoria scientifica della comprensione linguistica. Se l'esperienza fenomenologica non è una guida affidabile alla struttura dei processi cognitivi sottostanti, se la trasparenza apparente del significato maschera una complessità computazionale che non è direttamente accessibile all'introspezione cosciente, allora come possiamo sviluppare una teoria empiricamente adeguata e formalmente rigorosa di tali processi? Questo è precisamente il problema che ha afflitto la psicologia della comprensione linguistica fin dai suoi albori nel 19° secolo, quando i primi psicologi sperimentali come Wilhelm Wundt tentarono di applicare il metodo introspettivo allo studio dei processi mentali superiori e si scontrarono con l'opacità fenomenologica del pensiero e del linguaggio.

La risposta che la psicolinguistica contemporanea ha sviluppato è che dobbiamo affidarci necessariamente a metodi indiretti di indagine, a tecniche sperimentali che ci permettono di inferire la struttura dei processi cognitivi non-osservabili dalla loro firma comportamentale e neurale osservabile. Questi metodi includono: esperimenti comportamentali cronometrici che misurano tempi di reazione con precisione del millisecondo, permettendoci di costruire curve di latenza che riflettono la complessità computazionale relativa di diversi processi; analisi sistematiche di pattern di errori di produzione e comprensione, che possono rivelare i punti critici dove il sistema computazionale è più vulnerabile al breakdown; paradigmi di priming semantico, sintattico e morfologico, che sondano la struttura delle rappresentazioni mentali attraverso l'analisi degli effetti di facilitazione e interferenza; tecniche di neuroimaging funzionale come la risonanza magnetica funzionale (fMRI) e la magnetoencefalografia (MEG), che rivelano quali regioni cerebrali sono attive durante specifiche operazioni linguistiche e con quale dinamica spazio-temporale; e modelli computazionali implementati come simulazioni al computer, che generano predizioni quantitative testabili empiricamente attraverso confronti sistematici con dati comportamentali e neurali.

La Distinzione Fregeana tra Senso e Riferimento: Fondamenti della Semantica Moderna e Teoria della Composizionalità

La teoria semantica sviluppata da Gottlob Frege nel suo articolo seminale "Über Sinn und Bedeutung" del 1892 costituisce il punto di partenza assolutamente ineludibile per qualsiasi trattazione seria e sistematica della comprensione linguistica, e rappresenta uno dei contributi più profondi, influenti e duraturi alla filosofia del linguaggio, alla logica filosofica e alle fondazioni della semantica formale. L'intuizione fondamentale di Frege emerge da un puzzle logico-epistemologico profondo che riguarda la natura delle identità informative, un problema che aveva afflitto i logici e i filosofi almeno dai tempi di John Stuart Mill nel 19° secolo e che sembrava resistere a qualsiasi soluzione soddisfacente all'interno dei framework teorici disponibili all'epoca.

Consideriamo i seguenti due enunciati, che sono diventati l'esempio canonico e paradigmatico nella letteratura filosofica successiva:

(1) La stella del mattino è la stella del mattino (2) La stella del mattino è la stella della sera

Entrambi gli enunciati hanno la forma logica superficiale di un'identità, che possiamo rappresentare schematicamente come "a = b", dove a e b sono termini singolari che pretendono di riferirsi a oggetti specifici nel dominio del discorso. Tuttavia, nonostante questa identità formale nella struttura sintattica, questi due enunciati differiscono radicalmente e incontrovertibilmente nel loro valore cognitivo, o come Frege lo chiamava nel tedesco originale, nel loro Erkenntniswert—letteralmente, il loro "valore conoscitivo" o "valore epistemico".

L'enunciato (1) è una tautologia analitica, un'istanza del principio logico di identità (∀x: x = x), una verità logica che può essere conosciuta a priori, completamente indipendente da qualsiasi indagine empirica sul mondo fisico o astronomico. Qualsiasi parlante razionale e competente che comprende i termini "la stella del mattino" e l'operatore di identità "è" può riconoscere immediatamente, senza bisogno di osservazione empirica, che (1) deve essere vero, puramente in virtù della comprensione dei significati delle espressioni coinvolte e delle leggi della logica. Questo enunciato non fornisce alcuna informazione sostantiva, genuinamente estensionale, sul mondo; non ci dice nulla che non sapessimo già una volta che comprendiamo i termini che lo compongono. È, per usare la celebre terminologia introdotta da Immanuel Kant nella "Critica della Ragion Pura" del 1781—una distinzione che Frege conosceva intimamente e che giocava un ruolo centrale nel suo pensiero—analitico piuttosto che sintetico, una verità la cui negazione implicherebbe una contraddizione logica.

L'enunciato (2), al contrario, esprime una scoperta astronomica empirica, genuinamente sintetica, che le antiche civiltà mesopotamiche compirono attraverso osservazioni sistematiche e pazienti del cielo notturno nel corso di molti secoli. I Babilonesi, osservando con cura il movimento degli oggetti celesti, notarono che quello che appariva fenomenicamente come due oggetti celesti distinti e separati—un corpo luminoso particolarmente brillante visibile nel cielo orientale all'alba, poco prima del sorgere del sole (che chiamavano "la stella del mattino"), e un corpo luminoso ugualmente brillante visibile nel cielo occidentale al crepuscolo, poco dopo il tramonto (che chiamavano "la stella della sera")—era in realtà un unico pianeta, quello che nella nomenclatura astronomica moderna chiamiamo Venere, il secondo pianeta del sistema solare in ordine di distanza dal Sole. Questa scoperta, questa identità astronomica, non poteva essere determinata semplicemente attraverso l'analisi concettuale, la riflessione a priori o la pura logica; richiedeva necessariamente osservazione empirica prolungata, raccolta sistematica di dati osservazionali, e inferenza induttiva basata su pattern riconosciuti nel movimento celeste. L'enunciato (2) è quindi informativamente ricco, epistemicamente sostantivo, in un modo in cui (1) manifestamente non lo è.

Questo puzzle genera un problema teorico profondo. Se identifichiamo il significato (Bedeutung nella terminologia tecnica fregeana originale) di un'espressione semplicemente e direttamente con il suo riferimento—l'oggetto concreto nel mondo fisico a cui l'espressione si riferisce—allora non possiamo in alcun modo spiegare questa differenza cruciale nel valore cognitivo tra (1) e (2), poiché "la stella del mattino" e "la stella della sera" hanno manifestamente lo stesso riferimento, vale a dire il pianeta Venere, un corpo fisico sferico con un diametro di circa 12.104 chilometri, composto principalmente di roccia e metallo, che orbita attorno al Sole a una distanza media di circa 108 milioni di chilometri. Se il significato di un nome proprio, di una descrizione definita o di qualsiasi termine singolare è semplicemente e soltanto il suo portatore, l'oggetto che denota, allora i due termini "la stella

del mattino" e "la stella della sera" dovrebbero essere perfettamente sinonimi, completamente intercambiabili salva veritate (preservando il valore di verità) in tutti i contesti linguistici possibili, e non ci dovrebbe essere alcuna differenza informativa, alcuna differenza nel contenuto epistemico, tra gli enunciati (1) e (2). Ma questa conclusione è manifestamente, intuitivamente, incontrovertibilmente falsa: (1) è una tautologia banale e triviale mentre (2) esprime una scoperta scientifica sostanziale che ha richiesto secoli di osservazioni astronomiche per essere stabilita.

Frege concluse che dobbiamo distinguere, per ogni espressione significativa del linguaggio, due livelli semantici fondamentalmente distinti e irriducibili l'uno all'altro:

Il **Riferimento** (Bedeutung nel tedesco originale, anche se questo termine è notoriamente difficile da tradurre con precisione): l'oggetto concreto nel mondo a cui l'espressione si riferisce, il portatore del nome nel caso di nomi propri, l'estensione nel caso di predicati. Per le descrizioni definite "la stella del mattino" e "la stella della sera", il riferimento è numericamente identico: il pianeta Venere, un corpo fisico particolare che occupa una regione specifica dello spazio-tempo.

Il **Senso** (Sinn): il modo di presentazione (Art des Gegebenseins) del riferimento, la maniera specifica in cui l'espressione determina, seleziona o presenta il suo riferimento, il percorso cognitivo attraverso cui ci orientiamo mentalmente verso l'oggetto, la prospettiva concettuale sotto cui l'oggetto ci è dato all'apprensione cognitiva. "La stella del mattino" presenta Venere sotto il modo di presentazione: l'oggetto celeste particolarmente brillante visibile nel cielo orientale poco prima dell'alba; mentre "la stella della sera" presenta lo stesso pianeta sotto un modo di presentazione diverso: l'oggetto celeste brillante visibile nel cielo occidentale poco dopo il tramonto. Questi sono modi radicalmente diversi di pensare allo stesso oggetto, prospettive cognitive irriducibilmente distinte sullo stesso riferimento fisico.

La distinzione può essere formulata più precisamente nel seguente modo. Il senso di un'espressione è una funzione—nel senso matematico tecnico del termine—dal dominio dei contesti possibili (o delle "circostanze di valutazione" nella terminologia della semantica dei mondi possibili) al dominio delle estensioni. Formalmente, se E è un'espressione linguistica, possiamo rappresentare:

$$\text{Senso}(E) = \lambda c \,.\, \text{Riferimento}_c(E)$$

dove c varia su contesti o circostanze di valutazione (che potrebbero essere caratterizzati come coppie $\langle w,t \rangle$ di mondo possibile w e tempo t), e $\text{Riferimento}_c(E)$ è il riferimento che E ha nel contesto c. Il senso è quindi una funzione intensionale, mentre il riferimento è un'estensione particolare determinata da quella funzione applicata a un contesto specifico.

Per rendere questa astrazione più concreta e percepibile, consideriamo come la distinzione si applica sistematicamente attraverso le diverse categorie sintattiche fondamentali delle espressioni linguistiche:

Per i nomi propri e le descrizioni definite:

- Senso: il modo di presentazione dell'oggetto, il criterio cognitivo di identificazione, il "file" mentale associato al nome

- Riferimento: l'oggetto concreto denotato nel mondo, il portatore del nome

Per esempio, "Aristotele" e "il maestro di Alessandro Magno" hanno lo stesso riferimento (il filosofo greco vissuto nel 4° secolo a.C.) ma sensi diversi: il primo presenta l'individuo attraverso il suo nome proprio, il secondo attraverso una sua proprietà relazionale distintiva.

Per i predicati monadici (proprietà):

- Senso: il concetto (Begriff nella terminologia fregeana), inteso come una funzione da oggetti a valori di verità, una regola che specifica le condizioni necessarie e sufficienti che un oggetto deve soddisfare per cadere sotto il predicato
- Riferimento: l'estensione del predicato, ovvero l'insieme degli oggetti che soddisfano il predicato

Formalmente, se P è un predicato, possiamo rappresentare:

- Senso(P) = λx . [x soddisfa le condizioni definite da P]
- Riferimento(P) = {x | Senso(P)(x) = Vero}

Per esempio, per il predicato "è rosso":

- Senso: la funzione che mappa ogni oggetto x al valore Vero se x possiede la proprietà di rossezza, Falso altrimenti
- Riferimento: l'insieme di tutti gli oggetti rossi attualmente esistenti nel mondo

Per le frasi complete (enunciati dichiarativi):

- Senso: il pensiero (Gedanke) espresso, una proposizione in senso fregeano, un contenuto proposizionale che può essere vero o falso ma che trascende qualsiasi formulazione linguistica particolare
- Riferimento: il valore di verità, das Wahre (il Vero) o das Falsche (il Falso), considerati come oggetti astratti

Questa ultima tesi—che il riferimento di una frase completa sia semplicemente e soltanto il suo valore di verità—è forse l'aspetto più audace, contro-intuitivo e filosoficamente controverso della teoria fregeana. Essa implica che tutte le frasi vere, indipendentemente dal loro contenuto specifico, dalla loro struttura interna, dal dominio di realtà di cui parlano, hanno esattamente lo stesso riferimento (il Vero considerato come un oggetto astratto); e similmente, tutte le frasi false hanno lo stesso riferimento (il Falso). A prima vista, questo sembra assurdo e chiaramente inaccettabile: "la neve è bianca", "Roma è la capitale dell'Italia", "2+2=4" e "i protoni hanno carica positiva" sono frasi che parlano di domini completamente diversi della realtà—proprietà fisiche macroscopiche, relazioni geo-politiche, verità matematiche, e proprietà di particelle subatomiche rispettivamente—eppure secondo Frege tutte queste frasi avrebbero esattamente lo stesso riferimento semplicemente perché sono tutte vere. Come può una teoria che identifica il riferimento di frasi così radicalmente diverse essere adeguata?

La risposta di Frege è profonda e rivela la vera motivazione teorica sottostante alla sua proposta. Questa conseguenza apparentemente paradossale è in realtà essenziale, logicamente necessaria, per il progetto fondazionale fregeano di fornire una semantica

sistematicamente e rigorosamente compositiva per il linguaggio naturale e per i linguaggi logici formali. Il Principio di Composizionalità del Riferimento, che sta al cuore assoluto della teoria fregeana e che è forse il suo contributo teorico più duraturo e influente, può essere enunciato nel modo seguente:

Principio di Composizionalità Fregeano: Il riferimento di un'espressione complessa è una funzione deterministica del riferimento delle sue parti costituenti immediate e del modo specifico in cui queste parti sono sintatticamente combinate secondo le regole della grammatica.

Possiamo formalizzare questo principio nel modo seguente. Sia $[[\cdot]]$ la funzione di interpretazione semantica che assegna riferimenti alle espressioni linguistiche, mappando espressioni sintattiche in entità semantiche del tipo appropriato (oggetti, funzioni da oggetti a valori di verità, valori di verità, etc.). Allora, per un'espressione complessa E costruita dalle sottoespressioni immediate E_1, E_2, ..., E_n mediante l'operazione sintattica $\oplus$ (che potrebbe essere concatenazione, applicazione funzionale, modificazione, etc.), abbiamo il requisito fondamentale:

$$[[E_1 \oplus E_2 \oplus ... \oplus E_n]] = f_{\oplus}([[E_1]], [[E_2]], ..., [[E_n]])$$

dove $f_{\oplus}$ è una funzione semantica univocamente determinata dalla natura dell'operazione sintattica $\oplus$. Questa funzione prende come argomenti i riferimenti (le estensioni) delle parti costituenti immediate e restituisce come valore il riferimento (l'estensione) dell'espressione complessa risultante.

Ora, perché questo principio richiede che il riferimento di una frase sia il suo valore di verità? La ragione è la seguente. Consideriamo frasi complesse costruite attraverso operatori vero-funzionali come la negazione, la congiunzione e la disgiunzione. Per esempio, data la frase "piove", possiamo costruire la frase complessa "non piove", "piove e nevica", "piove o fa freddo", e così via. Il valore di verità di queste frasi complesse dipende esclusivamente dai valori di verità delle frasi componenti secondo le tavole di verità standard della logica proposizionale classica:

- [["non φ"]] = Vero se e solo se [[φ]] = Falso
- [["φ e ψ"]] = Vero se e solo se [[φ]] = Vero e [[ψ]] = Vero
- [["φ o ψ"]] = Vero se e solo se [[φ]] = Vero o [[ψ]] = Vero

Queste regole semantiche sono puramente vero-funzionali: determinano il valore di verità della frase complessa esclusivamente in base ai valori di verità delle frasi componenti, senza fare riferimento ad alcun aspetto della struttura interna di quelle frasi oltre al loro valore di verità. Ma se accettiamo il Principio di Composizionalità nella sua forma forte—che il riferimento di un'espressione complessa dipende soltanto dal riferimento delle sue parti—allora siamo costretti a concludere che il riferimento di una frase deve essere il suo valore di verità, poiché altrimenti non potremmo spiegare come le funzioni vero-funzionali operano compositivamente.

Per rendere questo argomento ancora più preciso, possiamo formularlo come segue. Supponiamo per assurdo che il riferimento di una frase φ non sia il suo valore di verità, ma qualche altra entità più ricca—call it Ref(φ). Consideriamo ora due frasi φ e ψ che hanno lo

stesso valore di verità (entrambe vere, diciamo) ma riferimenti diversi secondo questa ipotesi: Ref(φ) ≠ Ref(ψ). Ora formiamo la congiunzione "φ e χ" e "ψ e χ" per qualche terza frase χ. Se la congiunzione opera compositivamente sul riferimento delle sue parti, allora dovremmo avere:

[["φ e χ"]] = f_∧(Ref(φ), Ref(χ)) [["ψ e χ"]] = f_∧(Ref(ψ), Ref(χ))

Ma poiché Ref(φ) ≠ Ref(ψ) per ipotesi, e f_∧ è una funzione deterministica, avremmo [["φ e χ"]] ≠ [["ψ e χ"]]. Tuttavia, sappiamo dalla logica proposizionale che se φ e ψ hanno lo stesso valore di verità, allora "φ e χ" e "ψ e χ" devono avere lo stesso valore di verità. Quindi il riferimento della congiunzione non può essere determinato compositivamente dai riferimenti (nel senso non-vero-funzionale) delle sue parti. L'unico modo per preservare la composizionalità è identificare il riferimento di una frase con il suo valore di verità.

Questo ragionamento rivela che l'identificazione fregeana del riferimento di una frase con il suo valore di verità non è una bizzarria arbitraria o una stipulazione capricciosa, ma è richiesta necessariamente dalla combinazione di due tesi teoriche fondamentali: (1) il Principio di Composizionalità e (2) la natura vero-funzionale dei connettivi logici nel linguaggio naturale. È il prezzo che dobbiamo pagare se vogliamo una semantica compositiva rigorosa.

Consideriamo ora come questo framework teorico si applica concretamente all'analisi semantica di frasi semplici. Prendiamo l'esempio paradigmatico "Socrate è mortale". Secondo l'analisi fregeana, questa frase ha la seguente struttura semantica:

- [["Socrate"]] = Socrate (l'oggetto: il filosofo greco particolare che visse ad Atene nel 5°-4° secolo a.C.)
- [["è mortale"]] = λx . Mortale(x) (una funzione da oggetti a valori di verità: la funzione che mappa ogni oggetto x al valore Vero se x è mortale, Falso altrimenti)
- [["Socrate è mortale"]] = ["è mortale"] = (λx . Mortale(x))(Socrate) = Mortale(Socrate) = Vero

L'ultima uguaglianza vale perché Socrate è effettivamente mortale (è morto nel 399 a.C. dopo aver bevuto la cicuta). Questa derivazione mostra esattamente come il riferimento della frase complessa è determinato compositivamente dal riferimento del nome proprio "Socrate" (che è l'oggetto Socrate) e dal riferimento del predicato "è mortale" (che è una funzione da oggetti a valori di verità), attraverso la regola semantica dell'applicazione funzionale.

Per frasi più complesse che coinvolgono quantificazione, la situazione diventa considerevolmente più sottile e tecnicamente impegnativa. Consideriamo l'enunciato:

(3) Ogni uomo è mortale

Frege analizzò i quantificatori come predicati di secondo livello, ovvero predicati che si applicano non a oggetti individuali ma a concetti (che sono essi stessi funzioni da oggetti a valori di verità). La struttura semantica di (3) secondo Frege sarebbe:

[["ogni"]] = λF . λG . ∀x(F(x) → G(x))

Questa è una funzione che prende due concetti F e G come argomenti e restituisce il valore Vero se e solo se ogni oggetto che cade sotto il concetto F cade anche sotto il concetto G. Applicando questa analisi a (3):

[["uomo"]] = λx . Uomo(x) [["mortale"]] = λx . Mortale(x) [["ogni uomo"]] = ["ogni"] = λG . ∀x(Uomo(x) → G(x)) [["ogni uomo è mortale"]] = ["ogni uomo"] = ∀x(Uomo(x) → Mortale(x))

Questo è il valore Vero se e solo se ogni oggetto che è un uomo è anche mortale, che è precisamente il contenuto intuitivo della frase originale.

Il Principio di Composizionalità non si applica soltanto al riferimento ma anche, crucialmente, al senso. Frege formulò quello che possiamo chiamare il Principio di Composizionalità del Senso:

Principio di Composizionalità del Senso: Il senso di un'espressione complessa è una funzione del senso delle sue parti costituenti e del modo in cui sono combinate sintatticamente.

Formalmente, se Senso(·) è la funzione che assegna sensi alle espressioni, allora per un'espressione complessa $E_1 \oplus E_2 \oplus ... \oplus E_n$:

$Senso(E_1 \oplus E_2 \oplus ... \oplus E_n) = g_{\oplus}(Senso(E_1), Senso(E_2), ..., Senso(E_n))$

dove $g_{\oplus}$ è una funzione appropriata di combinazione di sensi determinata dall'operazione sintattica $\oplus$.

Questo principio è ciò che rende possibile la produttività illimitata del linguaggio naturale, quella proprietà straordinaria e distintiva dei linguaggi umani in virtù della quale possiamo comprendere e produrre un numero potenzialmente infinito di frasi nuove, mai incontrate prima nell'esperienza linguistica individuale. Possiamo comprendere infinite frasi nuove perché, una volta che abbiamo acquisito: (a) il lessico mentale che specifica il senso (e il riferimento) delle parole atomiche, e (b) le regole grammaticali che specificano come le parole possono essere combinate sintatticamente e come i sensi delle parti si compongono per formare il senso del tutto, possiamo determinare meccanicamente, attraverso un processo computazionale sistematico e algoritmico, il senso (e quindi il riferimento) di qualsiasi espressione complessa costruita da quelle parti secondo quelle regole combinatorie.

La teoria fregeana introduce anche una distinzione fondamentale tra ciò che Frege chiamava "pensiero" (Gedanke) e "rappresentazione" (Vorstellung). Il pensiero espresso da una frase—il suo senso—è oggettivo, pubblicamente accessibile, invariante attraverso diversi parlanti e diversi tempi. Quando io penso che la neve è bianca e tu pensi che la neve è bianca, cogliamo lo stesso pensiero, la stessa entità proposizionale astratta. Le rappresentazioni mentali, al contrario, sono soggettive, private, potenzialmente diverse per ogni individuo. La mia immagine mentale della neve potrebbe essere molto diversa dalla tua, ma il pensiero che cogliamo è identico. Formalmente:

Pensiero(φ) = Senso(φ) [oggettivo, pubblico, intersoggettivamente condivisibile] Rappresentazione_i(φ) = la particolare realizzazione psicologica del pensiero nella mente dell'individuo i [soggettivo, privato, idiosincratico]

Questa distinzione è cruciale per la teoria fregeana della comunicazione. Quando un parlante A enuncia una frase φ e un ascoltatore B la comprende, ciò che viene trasmesso non è la rappresentazione psicologica privata di A (che sarebbe inaccessibile a B), ma il pensiero oggettivo espresso da φ, che può essere colto da entrambi A e B in virtù della loro competenza linguistica condivisa.

La teoria fregeana solleva però numerosi problemi teorici profondi che hanno generato decenni di dibattito filosofico. Uno dei problemi più seri riguarda i nomi propri privi di portatore. Consideriamo il nome "Sherlock Holmes". Secondo l'analisi fregeana standard, questo nome ha un senso (qualcosa come: il celebre detective che vive al 221B di Baker Street, esperto di deduzione, creato da Arthur Conan Doyle) ma nessun riferimento, poiché Sherlock Holmes non esiste realmente. Ma allora cosa dire di frasi come "Sherlock Holmes vive a Londra"? Se "Sherlock Holmes" manca di riferimento, e il riferimento della frase dipende compositivamente dal riferimento delle sue parti, allora la frase dovrebbe mancare di valore di verità—non sarebbe né vera né falsa. Questo sembra contro-intuitivo: vogliamo dire che la frase è vera all'interno del contesto finzionale delle storie di Conan Doyle.

Frege stesso era consapevole di questo problema e propose che nel contesto della finzione letteraria, i nomi propri esprimono un senso ma non pretendono di avere un riferimento; la questione del riferimento è, per così dire, sospesa. Altri filosofi hanno proposto soluzioni alternative: alcuni sostengono che "Sherlock Holmes" si riferisce a un oggetto astratto (un personaggio finzionale); altri propongono analisi in termini di mondi possibili o situazioni finzionali; altri ancora abbracciano la conseguenza che frasi finzionali sono genuinamente prive di valore di verità nel senso standard e introducono operatori finzionali che modificano le condizioni di verità.

Un secondo problema profondo riguarda i contesti opachi o intensionali—contesti in cui la sostituzione di coreferenziali non preserva il valore di verità. Consideriamo:

(4) Giovanni crede che la stella del mattino sia visibile all'alba (5) Giovanni crede che la stella della sera sia visibile all'alba

Anche se "la stella del mattino" e "la stella della sera" hanno lo stesso riferimento (Venere), è perfettamente possibile che (4) sia vera mentre (5) è falsa—Giovanni potrebbe non sapere che la stella del mattino è la stella della sera, e quindi potrebbe credere che la prima sia visibile all'alba ma non credere questo della seconda (o addirittura credere esplicitamente il contrario).

Questo viola il Principio di Composizionalità del Riferimento nella sua forma naive: il valore di verità di una frase non è sempre determinato esclusivamente dal riferimento delle sue parti, poiché la sostituzione di "la stella del mattino" con "la stella della sera" (che hanno lo stesso riferimento) in (4) può cambiare il valore di verità. Frege propose una soluzione ingegnosa: nei contesti di credenza e altri contesti intensionali, le espressioni non hanno il loro riferimento abituale ma hanno invece come riferimento ciò che normalmente sarebbe il loro senso. In altre parole:

Riferimento_usuale("la stella del mattino") = Venere Riferimento_in_contesto_di_credenza("la stella del mattino") = Senso_usuale("la stella del mattino") = il modo di presentazione: l'oggetto brillante visibile all'alba

Questo fenomeno, che Frege chiamava "riferimento indiretto" (ungerade Bedeutung), preserva la composizionalità ma al prezzo di complicare considerevolmente la teoria semantica, richiedendo una gerarchia infinita di sensi e riferimenti (senso di primo livello, senso di secondo livello quando si parla di credenze su credenze, etc.).

La Teoria delle Descrizioni di Russell: Forma Logica versus Forma Grammaticale e l'Analisi Quantificazionale delle Descrizioni Definite

Bertrand Russell, nell'articolo monumentale "On Denoting" del 1905, estese e raffinò profondamente l'analisi fregeana affrontando un puzzle semantico che aveva afflitto Frege stesso e che sembrava minacciare le fondazioni stesse della teoria semantica compositiva: il problema delle descrizioni definite apparentemente non denotanti, ovvero quelle espressioni della forma "il F" (dove F è un predicato) che grammaticalmente sembrano riferirsi a un oggetto specifico ma per le quali non esiste alcun oggetto corrispondente nella realtà.

Consideriamo l'enunciato che è diventato l'esempio paradigmatico nella letteratura filosofica:

(6) L'attuale re di Francia è calvo

Questo enunciato sembra perfettamente significativo dal punto di vista grammaticale—è una frase ben formata dell'italiano che segue tutte le regole sintattiche standard—eppure presenta un problema semantico profondo e apparentemente intrattabile: non esiste attualmente alcun re di Francia, poiché la Francia è una repubblica dal 1870 (con una breve interruzione durante il regime di Vichy). La Francia non ha un monarca regnante.

Se trattiamo "l'attuale re di Francia" come un termine singolare referenziale genuino che deve riferirsi a qualche oggetto specifico per poter contribuire semanticamente al valore di verità della frase contenente, ci troviamo in una situazione teorica estremamente difficile. Tre posizioni logicamente possibili si presentano:

Posizione 1 (Meinongiana): La frase (6) è vera o falsa (ha un valore di verità determinato), e "l'attuale re di Francia" si riferisce effettivamente a qualcosa, ma questo qualcosa è un oggetto che sussiste (besteht) senza esistere (existieren). Questa posizione, difesa vigorosamente dal filosofo austriaco Alexius Meinong nei primi anni del 20° secolo, accetta un'ontologia radicalmente inflazionata che include non solo oggetti esistenti nel mondo fisico attuale ma anche oggetti meramente possibili, oggetti impossibili (come il quadrato rotondo), oggetti incompleti, e così via. Secondo Meinong, "l'attuale re di Francia" denota un oggetto che ha la proprietà di essere re di Francia e di essere attuale, ma che semplicemente non ha la proprietà (di ordine superiore) di esistenza.

Posizione 2 (Fregeana): La frase (6) non ha alcun valore di verità—non è né vera né falsa—perché una delle sue componenti costituenti (il termine singolare "l'attuale re di Francia") manca di riferimento. Quando una componente di una frase complessa manca di riferimento, l'intera frase manca di riferimento, e quindi manca di valore di verità. Secondo il principio compositivo fregeano, se [[t]] non è definito per qualche termine t che occorre in φ,

allora [[φ]] non è definito. Questa è la posizione che Frege stesso abbracciò, ma con una certa riluttanza, poiché complica la logica introducendo gap di valore di verità che violano il principio classico di bivalenza (secondo cui ogni enunciato dichiarativo è o vero o falso).

Posizione 3 (Russelliana): La frase (6) è significativa e ha un valore di verità determinato—specificamente, è falsa—ma "l'attuale re di Francia" non è genuinamente un termine singolare referenziale nonostante la sua apparenza grammaticale superficiale. La forma grammaticale superficiale dell'enunciato maschera la sua vera forma logica sottostante, che è molto diversa da ciò che la sintassi superficiale suggerirebbe.

Russell adottò decisamente la terza posizione, proponendo la sua celebre e rivoluzionaria teoria delle descrizioni definite, che ha avuto un'influenza profonda e duratura non solo sulla filosofia del linguaggio ma anche sulla linguistica teorica, sulla logica filosofica e sulle fondazioni della semantica formale. Secondo questa teoria, le descrizioni definite sono espressioni quantificazionali in disguise, maschere grammaticali di strutture logiche più profonde. La forma grammaticale superficiale di (6) maschera radicalmente la sua vera forma logica, che può essere rivelata solo attraverso un'analisi logica appropriata.

Russell propose che (6) dovrebbe essere analizzato come avente la seguente forma logica:

(6') $\exists x$ [Re-attuale-di-Francia(x) $\wedge$ $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x) $\wedge$ Calvo(x)]

In italiano esplicito: "Esiste un x tale che x è attualmente re di Francia, e per ogni y, se y è attualmente re di Francia allora y è identico a x, e x è calvo".

Questa analisi scompone la descrizione definita "l'attuale re di Francia" in tre componenti logiche distinte e separabili:

1. **Componente Esistenziale:** $\exists x$ Re-attuale-di-Francia(x) Questa componente afferma che esiste almeno un oggetto che è attualmente re di Francia. In simboli logici: l'insieme {x | Re-attuale-di-Francia(x)} è non-vuoto.

2. **Componente di Unicità:** $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x) Questa componente afferma che esiste al massimo un oggetto che è attualmente re di Francia, ovvero che se qualche oggetto y ha questa proprietà, allora y deve essere identico a x (l'oggetto la cui esistenza è affermata dalla componente esistenziale). In simboli: |{x | Re-attuale-di-Francia(x)}| ≤ 1.

3. **Componente Predicativa:** Calvo(x) Questa componente attribuisce la proprietà di essere calvo all'unico oggetto (se esiste) che soddisfa la descrizione.

Quando una qualsiasi di queste tre componenti è falsa, l'intera congiunzione è falsa, e quindi la frase originale (6) è falsa. Nel caso attuale, la componente esistenziale è manifestamente falsa (non esiste alcun re attuale di Francia nel mondo reale del 2024), quindi la formula (6') nel suo insieme è falsa. Di conseguenza, l'enunciato originale (6) è falso, non privo di valore di verità come nella soluzione fregeana, e non vero in virtù del riferimento a un oggetto meinongiano non-esistente.

L'importanza teorica di questa analisi va immensamente al di là della risoluzione di un puzzle tecnico apparentemente marginale sulle descrizioni definite non denotanti. Russell dimostrò un principio metodologico di portata generale che è diventato centrale nella filosofia analitica del 20° secolo e nella linguistica teorica moderna: la forma grammaticale superficiale di un enunciato può divergere radicalmente, e in modi sistematici e profondi, dalla sua vera forma logica sottostante, e solo attraverso un'analisi logica rigorosa possiamo rivelare la struttura semantica profonda che determina il comportamento inferenziale e le condizioni di verità dell'enunciato. La grammatica superficiale può ingannare; la logica rivela la verità.

Possiamo generalizzare l'analisi russelliana a qualsiasi descrizione definita della forma "il F" (dove F è un predicato arbitrario). La descrizione "il F" contribuisce alla forma logica di una frase "il F è G" attraverso la seguente trasformazione:

"il F è G" $\leftrightarrow \exists x\,[F(x) \wedge \forall y(F(y) \rightarrow y = x) \wedge G(x)]$

Questa può essere letta come: "Esiste esattamente un F, e quel F è G". La clausola "esiste esattamente un" è analizzata come la congiunzione di esistenza ($\exists x\ F(x)$) e unicità ($\forall y(F(y) \rightarrow y = x)$).

Possiamo introdurre un operatore di descrizione definita ι (la iota di Russell) per rendere questa analisi più compatta e trasparente. Definiamo:

$\iota x\ F(x)$ = l'unico x tale che F(x)

con la seguente clausola di eliminazione:

$G(\iota x\ F(x)) \leftrightarrow \exists x\,[F(x) \wedge \forall y(F(y) \rightarrow y = x) \wedge G(x)]$

Questa notazione ci permette di scrivere (6) più elegantemente come:

(6'') Calvo(ιx Re-attuale-di-Francia(x))

che si espande alla forma logica completa (6') attraverso l'applicazione della clausola di eliminazione.

Consideriamo ora come questa analisi cattura elegantemente le nostre intuizioni semantiche in casi meno problematici. Prendiamo l'esempio celebre:

(7) L'autore di Waverley era scozzese

Sir Walter Scott fu l'autore del romanzo "Waverley" pubblicato anonimamente nel 1814, e Scott era effettivamente scozzese. Secondo l'analisi russelliana, (7) ha la forma logica:

(7') $\exists x\,[\text{Autore-di-Waverley}(x) \wedge \forall y(\text{Autore-di-Waverley}(y) \rightarrow y = x) \wedge \text{Scozzese}(x)]$

Questa forma logica spiega perfettamente perché (7) è informativa e non-triviale anche se sappiamo già che Sir Walter Scott era scozzese. L'informazione che l'autore di Waverley è identico a Scott non è contenuta esplicitamente nella forma logica di (7); la frase parla di chiunque abbia scritto Waverley, caratterizzandolo attraverso questa relazione funzionale

con l'opera letteraria, senza menzionare Scott per nome proprio. La frase avrebbe lo stesso significato e le stesse condizioni di verità anche in un mondo possibile alternativo dove qualcun altro avesse scritto Waverley.

Russell utilizzò questa teoria per risolvere anche altri puzzle filosofici profondi. Consideriamo l'enunciato problematico:

(8) L'attuale re di Francia non esiste

Se trattiamo "l'attuale re di Francia" come un termine singolare referenziale genuino, questo enunciato sembra contraddittorio o paradossale: stiamo apparentemente dicendo di qualcosa (l'attuale re di Francia) che esso non esiste, ma come possiamo riferirci a qualcosa per poi negarne l'esistenza? Se "l'attuale re di Francia" si riferisce effettivamente a un oggetto, allora quell'oggetto esiste, e (8) è falso. Ma se non si riferisce a nulla, allora come può l'enunciato essere significativo?

L'analisi russelliana dissolve questo puzzle. Secondo Russell, (8) è ambiguo tra due letture diverse a seconda dello scope (ambito) della negazione rispetto alla descrizione definita:

Lettura con scope primario della descrizione (primary occurrence): (8a) $\exists x$ [Re-attuale-di-Francia(x) $\wedge$ $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x) $\wedge$ $\neg$Esiste(x)]

Questa lettura è contraddittoria: afferma che esiste un unico re attuale di Francia e che quell'individuo non esiste. Questa è ovviamente falsa in virtù della contraddizione logica.

Lettura con scope primario della negazione (secondary occurrence): (8b) $\neg \exists x$ [Re-attuale-di-Francia(x) $\wedge$ $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x)]

Questa lettura afferma semplicemente che non esiste un unico re attuale di Francia, ovvero nega l'esistenza di un tale individuo. Questa è vera, e cattura quello che normalmente intendiamo quando proferiamo (8).

Tuttavia, Russell notava che nella seconda lettura dovremmo propriamente scrivere la negazione dell'intera formula esistenziale, che è equivalente a:

(8b') $\neg \exists x$ [Re-attuale-di-Francia(x) $\wedge$ $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x)] $\equiv$ $\forall x$ $\neg$[Re-attuale-di-Francia(x) $\wedge$ $\forall y$(Re-attuale-di-Francia(y) $\rightarrow$ y = x)] $\equiv$ $\forall x$ [Re-attuale-di-Francia(x) $\rightarrow$ $\exists y$(Re-attuale-di-Francia(y) $\wedge$ y $\neq$ x)] $\vee$ $\forall x$ $\neg$Re-attuale-di-Francia(x)

Questa formula afferma che o non c'è nessun re di Francia, o ce n'è più di uno. In entrambi i casi, la descrizione definita "l'attuale re di Francia" non è soddisfatta.

La teoria delle descrizioni di Russell ha conseguenze filosofiche profonde per la natura della conoscenza per conoscenza diretta (knowledge by acquaintance) versus conoscenza per descrizione (knowledge by description). Russell sosteneva che possiamo avere conoscenza diretta solo di un numero molto ristretto di entità: i nostri dati sensoriali immediati, forse noi stessi, e possibilmente alcuni universali come il rossore o la rotondità. Tutto il resto della nostra conoscenza del mondo è mediato da descrizioni definite. Per esempio, non abbiamo conoscenza diretta di Napoleone Bonaparte, ma solo conoscenza per descrizione:

"l'imperatore francese che fu sconfitto a Waterloo nel 1815", "il generale che invase la Russia nel 1812", etc. Quando pensiamo a Napoleone, il contenuto del nostro pensiero è dato non da un riferimento diretto a Napoleone stesso (che è morto da molto tempo e con cui non abbiamo alcun contatto causale diretto), ma da qualche descrizione definita o insieme di descrizioni che "selezionano" Napoleone.

Questa distinzione epistemologica ha ramificazioni semantiche importanti. I veri nomi propri logici (logically proper names), secondo Russell, sono solo quelli che si riferiscono direttamente a entità con cui abbiamo conoscenza diretta. Questi sono termini come "questo" e "quello" usati per riferirsi a dati sensoriali presenti. Tutti gli altri nomi apparenti—"Napoleone", "Roma", "Venere"—sono in realtà descrizioni definite mascherate. "Napoleone" abbrevia qualcosa come "l'uomo chiamato 'Napoleone' che fece X, Y, Z". Questa è la teoria delle descrizioni nascoste (disguised descriptions).

La teoria russelliana può essere formalizzata più rigorosamente nel seguente modo. Introduciamo un linguaggio formale L_DD (il linguaggio delle descrizioni definite) con la seguente sintassi:

Termini:

- Costanti individuali: a, b, c, ...
- Variabili: x, y, z, ...
- Descrizioni definite: ιx φ(x), dove φ(x) è una formula con x libera

Formule atomiche:

- $P(t_1, ..., t_n)$, dove P è un predicato n-ario e $t_1, ..., t_n$ sono termini

Formule complesse:

- $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, $\forall x\, \varphi$, $\exists x\, \varphi$

La semantica è data da una struttura $M = \langle D, I \rangle$ dove D è un dominio non-vuoto e I è una funzione di interpretazione. Per le descrizioni definite, abbiamo la seguente clausola semantica:

Per ogni formula φ(t) dove t = ιx ψ(x), definiamo: $M \models \varphi(\iota x\, \psi(x))$ sse $M \models \exists x\, [\psi(x) \wedge \forall y(\psi(y) \rightarrow y = x) \wedge \varphi(x)]$

Questa clausola elimina le descrizioni definite riducendole a quantificatori esistenziali con condizioni di unicità.

Un aspetto sottile e tecnicamente importante della teoria russelliana riguarda le proprietà di scope delle descrizioni definite in frasi complesse. Consideriamo:

(9) Giovanni crede che l'autore di Waverley sia scozzese

Questa frase è ambigua tra due letture:

Lettura de dicto (scope ristretto della descrizione): (9a) Crede(Giovanni, [∃x [Autore-di-Waverley(x) ∧ ∀y(Autore-di-Waverley(y) → y = x) ∧ Scozzese(x)]])

Questa lettura afferma che Giovanni crede la proposizione che esiste un unico autore di Waverley e che quell'autore è scozzese. Questa è una credenza de dicto—una credenza su come sono le cose, descritte in un certo modo.

Lettura de re (scope largo della descrizione): (9b) ∃x [Autore-di-Waverley(x) ∧ ∀y(Autore-di-Waverley(y) → y = x) ∧ Crede(Giovanni, Scozzese(x))]

Questa lettura afferma che esiste un unico autore di Waverley, e di quell'individuo (che risulta essere Scott), Giovanni crede che sia scozzese. Questa è una credenza de re—una credenza su un particolare oggetto nel mondo.

La distinzione è cruciale perché le due letture possono avere valori di verità diversi. Supponiamo che Giovanni non sappia che Scott ha scritto Waverley. Allora (9a) potrebbe essere falsa (Giovanni non crede che l'autore di Waverley sia scozzese, perché non sa chi è l'autore), mentre (9b) potrebbe essere vera (dell'individuo Scott, che è di fatto l'autore di Waverley, Giovanni crede che sia scozzese).

Formalmente, possiamo rappresentare l'ambiguità di scope attraverso indicizzazione esplicita. Sia α[ιx φ(x) / n] la formula α con la descrizione ιx φ(x) marcata con l'indice di scope n. Le regole di interpretazione sono:

1. Se n è il scope più esterno in α, allora: ⟦α[ιx φ(x) / n]⟧ = ⟦∃x [φ(x) ∧ ∀y(φ(y) → y = x) ∧ α[x / n]]⟧

2. Se n non è il scope più esterno, continua a percolate verso l'esterno finché non lo diventa.

Russell difese anche la sua teoria contro varie obiezioni. Una obiezione classica, sollevata da P.F. Strawson nel 1950, è che la teoria di Russell confonde il significato con l'uso. Quando un parlante pronuncia (6) "L'attuale re di Francia è calvo", secondo Strawson, il parlante non fa un'asserzione falsa ma piuttosto fallisce nel fare un'asserzione del tutto, perché la presupposizione esistenziale (che esiste un re di Francia) non è soddisfatta. La descrizione "l'attuale re di Francia" non afferma ma presuppone l'esistenza del suo referente.

Possiamo formalizzare l'approccio di Strawson nel modo seguente. Una frase φ contenente una descrizione definita ιx ψ(x) ha:

- **Presupposizione:** ∃x [ψ(x) ∧ ∀y(ψ(y) → y = x)]
- **Asserzione:** Se la presupposizione è soddisfatta, allora φ è valutata per verità/falsità. Se la presupposizione non è soddisfatta, φ manca di valore di verità.

Formalmente, introduciamo una logica trivalente con valori di verità {Vero, Falso, Indefinito}:

⟦φ(ιx ψ(x))⟧ = { ⟦φ(a)⟧ se ∃!x ψ(x) e a è l'unico x tale che ψ(x) Indefinito se ¬∃!x ψ(x) }

Questo approccio cattura l'intuizione che (6) non è semplicemente falso ma è in qualche modo "difettoso" o "inadeguato" quando pronunciato in un contesto dove non c'è re di Francia.

Russell rispose che questa distinzione tra asserzione e presupposizione, sebbene possa catturare certe intuizioni pragmatiche, complica inutilmente la logica e la semantica. È più semplice e teoricamente più elegante trattare tutte le implicazioni esistenziali come parte del contenuto assertivo della frase, soggetto a valutazione vero-funzionale standard.

Semantica Vero-Condizionale: La Tradizione Davidson-Tarski e la Teoria Compositiva della Verità

La teoria vero-condizionale del significato, articolata nella sua forma più influente e sistematica da Donald Davidson negli anni '60 e '70 del 20° secolo, propone che comprendere il significato di un enunciato consista fondamentalmente nel conoscere le sue condizioni di verità—le condizioni che devono ottenere nel mondo affinché l'enunciato sia vero. Questa idea, nelle sue formulazioni preliminari, può essere fatta risalire almeno a Ludwig Wittgenstein nel "Tractatus Logico-Philosophicus" del 1921, dove scrisse la celebre tesi: "Comprendere una proposizione significa sapere che cosa è il caso se essa è vera" (4.024). Ma fu Davidson che sviluppò questa intuizione in una teoria semantica formale rigorosa e sistematica, ispirandosi profondamente al lavoro del logico e matematico polacco Alfred Tarski sulla definizione formale della verità.

Tarski, nel suo articolo monumentale "Der Wahrheitsbegriff in den formalisierten Sprachen" ("Il concetto di verità nei linguaggi formalizzati") del 1933, affrontò un problema che aveva afflitto logici e filosofi fin dall'antichità: come possiamo fornire una definizione rigorosa, matematicamente precisa, del concetto di verità che sia allo stesso tempo materialmente adeguata (che catturi correttamente le nostre intuizioni pre-teoriche su quando le frasi sono vere) e formalmente coerente (che non generi paradossi come il celebre paradosso del mentitore)?

Il paradosso del mentitore, nella sua formulazione classica, emerge dalla seguente frase autoeferenziale:

(L) Questa frase è falsa.

Se (L) è vera, allora ciò che afferma deve essere il caso, quindi (L) è falsa. Ma se (L) è falsa, allora ciò che afferma non è il caso, quindi (L) non è falsa, quindi è vera. In entrambi i casi arriviamo a una contraddizione: (L) è vera se e solo se è falsa. Questo mostra che una teoria naive della verità—che permette di affermare semplicemente che una frase è vera se e solo se le cose sono come la frase dice che sono—porta a contraddizione quando applicata a frasi autoreferenziali.

La soluzione di Tarski fu duplice. Primo, distinse rigorosamente tra il linguaggio oggetto (il linguaggio di cui stiamo fornendo una teoria semantica) e il metalinguaggio (il linguaggio in cui formuliamo la teoria semantica). La definizione di verità per il linguaggio oggetto deve essere data nel metalinguaggio, che deve essere semanticamente più ricco del linguaggio oggetto (deve contenere i mezzi per parlare della sintassi e della semantica del linguaggio oggetto). Questo blocca i paradossi autoreferenziali perché una frase del linguaggio oggetto

non può predicare la verità di se stessa—può essere solo il metalinguaggio a predicare verità o falsità delle frasi del linguaggio oggetto.

Secondo, Tarski fornì una definizione ricorsiva della verità, costruendo dalle condizioni di verità degli enunciati atomici (le formule più semplici) alle condizioni di verità degli enunciati complessi costruiti attraverso connettivi logici e quantificatori. Questo approccio composizionale è ciò che rende la definizione formalmente rigorosa e matematicamente trattabile.

Per un linguaggio formale L del primo ordine con un insieme specificato di formule ben formate, Tarski definì la nozione di verità in un modello (o struttura) M. Un modello è una coppia $M = \langle D, I \rangle$ dove:

- D è un dominio non-vuoto di oggetti (l'universo del discorso)
- I è una funzione di interpretazione che assegna significati semantici alle espressioni sintattiche di L:
    - Per ogni costante individuale c, $I(c) \in D$
    - Per ogni predicato monadico P, $I(P) \subseteq D$
    - Per ogni predicato diadico R, $I(R) \subseteq D \times D$
    - E così via per predicati n-adici: $I(P^n) \subseteq D^n$

La verità è definita relativamente a un modello M e a un'assegnazione variabile g (una funzione da variabili a elementi del dominio: $g: Var \rightarrow D$). Scriviamo $M, g \vDash \varphi$ per dire che la formula φ è vera nel modello M sotto l'assegnazione g.

Le clausole ricorsive della definizione tarskiana sono:

Formule atomiche: $M, g \vDash P(t_1, ..., t_n)$ sse $\langle [\![t_1]\!]^{M,g}, ..., [\![t_n]\!]^{M,g} \rangle \in I(P)$

dove $[\![t]\!]^{M,g}$ è il valore semantico (denotazione) del termine t nel modello M sotto l'assegnazione g:

- $[\![c]\!]^{M,g} = I(c)$ se c è una costante
- $[\![x]\!]^{M,g} = g(x)$ se x è una variabile

Negazione: $M, g \vDash \neg\varphi$ sse non $M, g \vDash \varphi$

Ovvero, ¬φ è vera sse φ non è vera.

Congiunzione: $M, g \vDash (\varphi \wedge \psi)$ sse $M, g \vDash \varphi$ e $M, g \vDash \psi$

Disgiunzione: $M, g \vDash (\varphi \vee \psi)$ sse $M, g \vDash \varphi$ oppure $M, g \vDash \psi$

Condizionale materiale: $M, g \vDash (\varphi \rightarrow \psi)$ sse non $M, g \vDash \varphi$ oppure $M, g \vDash \psi$

Equivalentemente: $M, g \vDash (\varphi \rightarrow \psi)$ sse se $M, g \vDash \varphi$ allora $M, g \vDash \psi$

Quantificatore esistenziale: $M, g \vDash \exists x\, \varphi(x)$ sse esiste $d \in D$ tale che $M, g[x/d] \vDash \varphi(x)$

dove g[x/d] è l'assegnazione identica a g eccetto che assegna d alla variabile x: g$\underline{x/d}$ = { d se y = x g(y) se y ≠ x }

Quantificatore universale: M, g ⊨ ∀x φ(x) sse per ogni d ∈ D, M, g[x/d] ⊨ φ(x)

Questa definizione ricorsiva permette di determinare meccanicamente, attraverso un processo computazionale sistematico, se una formula qualsiasi è vera in un modello dato, purché la formula sia costruita mediante un numero finito di applicazioni delle regole sintattiche del linguaggio. La bellezza di questa definizione sta nella sua natura compositiva: la verità di una formula complessa è determinata sistematicamente dalla verità delle sue sottoformule attraverso clausole ricorsive che rispecchiano esattamente la struttura sintattica del linguaggio.

Tarski dimostrò che questa definizione è materialmente adeguata nel senso seguente. Per ogni frase φ del linguaggio oggetto (una formula senza variabili libere), la teoria genera un teorema della forma:

φ è vera in L se e solo se p

dove φ è il nome della frase (ottenuto tipicamente mettendo la frase tra virgolette) e p è una frase del metalinguaggio che specifica le condizioni sotto cui φ è vera. Questi sono chiamati T-sentences o teoremi-T (T-biconditionals) in onore di Tarski.

Per esempio, se il linguaggio oggetto è l'italiano e il metalinguaggio è anch'esso l'italiano (ma arricchito con terminologia semantica), la teoria dovrebbe generare teoremi come:

- "La neve è bianca" è vera in italiano sse la neve è bianca
- "Roma è la capitale dell'Italia" è vera in italiano sse Roma è la capitale dell'Italia
- "Piove" è vera in italiano sse piove
- "∀x(Uomo(x) → Mortale(x))" è vera sse ogni uomo è mortale

Tarski impose come condizione di adeguatezza materiale (Convenzione T) che una definizione di verità per un linguaggio L sia adeguata solo se implica tutte le T-sentences corrette per le frasi di L. Un teorema-T è corretto se la frase p nel lato destro del biconditionale è una traduzione accurata della frase φ menzionata nel lato sinistro, ovvero se p e φ esprimono le stesse condizioni di verità.

Donald Davidson, fortemente influenzato dal lavoro di Tarski, propose negli anni '60 di estendere questo approccio dai linguaggi formali artificiali ai linguaggi naturali come l'italiano, l'inglese, il tedesco, etc. La sua idea centrale, radicale per l'epoca, era che fornire una teoria del significato per una lingua naturale L equivale a fornire una teoria della verità tarskiana per L. Una teoria del significato adeguata, secondo Davidson, deve essere una teoria compositiva che genera sistematicamente le condizioni di verità per ogni frase possibile della lingua a partire da un numero finito di assiomi semantici per le parole e le costruzioni sintattiche.

L'insight profondo di Davidson era che se una teoria della verità per L genera tutte le T-sentences corrette, e le genera in modo compositivo (ovvero derivando le condizioni di verità delle frasi complesse sistematicamente dalle condizioni di verità o dai contributi

semantici delle loro parti), allora tale teoria cattura implicitamente il significato delle espressioni di L. La ragione è che conoscere le condizioni di verità di una frase è già, secondo Davidson, conoscere il suo significato. Se sappiamo che "La neve è bianca" è vera se e solo se la neve è bianca, allora sappiamo cosa significa quella frase—sappiamo che cosa deve essere il caso affinché la frase sia vera.

Ma Davidson non si limitò a richiedere che la teoria generasse T-sentences corretti. Richiese anche—e questo è cruciale—che la teoria fosse compositiva, ovvero che derivasse le condizioni di verità delle frasi complesse sistematicamente dalle condizioni di verità (o meglio, dai contributi semantici) delle loro parti costituenti. Questo requisito di composizionalità è ciò che distingue una genuina teoria del significato da un mero elenco o catalogo di biconditioni vere ma teoricamente inerte. Una teoria che semplicemente elencasse:

- "La neve è bianca" è vera sse la neve è bianca
- "L'erba è verde" è vera sse l'erba è verde
- "Roma è la capitale dell'Italia" è vera sse Roma è la capitale dell'Italia
- ...

per ogni frase della lingua, sarebbe materialmente adeguata (tutte le T-sentences sarebbero corrette) ma sarebbe teoricamente inutile perché non catturerebbe le regolarità sistematiche che permettono ai parlanti di comprendere infinite frasi nuove. Una teoria compositiva, al contrario, deriva queste T-sentences da un numero finito di assiomi che specificano il contributo semantico delle parole individuali e delle regole di combinazione sintattica.

Per illustrare concretamente come una teoria davidsoniana potrebbe funzionare, consideriamo come potrebbe trattare un frammento semplificato dell'italiano. Abbiamo bisogno di assiomi semantici per il lessico e regole compositive per le costruzioni sintattiche.

Assiomi lessicali per termini individuali (nomi propri):

- Riferimento("Roma") = Roma
- Riferimento("Maria") = Maria
- Riferimento("Giovanni") = Giovanni
- Riferimento("Socrate") = Socrate

Questi assiomi specificano che il riferimento del nome "Roma" è la città di Roma, il riferimento di "Maria" è l'individuo Maria, e così via.

Assiomi lessicali per predicati monadici:

- Per ogni oggetto x, x soddisfa "città" sse x è una città
- Per ogni oggetto x, x soddisfa "corre" sse x corre
- Per ogni oggetto x, x soddisfa "mortale" sse x è mortale
- Per ogni oggetto x, x soddisfa "filosofo" sse x è un filosofo

Formalmente: Soddisfa(x, "P") sse P(x), dove P è il predicato corrispondente nel metalinguaggio.

Assiomi lessicali per predicati diadici (relazioni):

- Per ogni coppia di oggetti x, y: ⟨x,y⟩ soddisfa "ama" sse x ama y
- Per ogni coppia di oggetti x, y: ⟨x,y⟩ soddisfa "vede" sse x vede y

Regole compositive per combinazioni soggetto-predicato:

Per una frase della forma [NP VP], dove NP è un sintagma nominale e VP è un sintagma verbale:

Regola 1: Se NP = t (un termine singolare) e VP = "è P" (dove P è un predicato), allora: "t è P" è vera sse Soddisfa(Riferimento(t), "P")

Applicando questa regola, possiamo derivare le seguenti T-sentences:

- "Roma è una città" è vera sse Soddisfa(Riferimento("Roma"), "città")
  sse Soddisfa(Roma, "città") sse Roma è una città

- "Maria corre" è vera sse Soddisfa(Riferimento("Maria"), "corre") sse Soddisfa(Maria, "corre")
  sse Maria corre

- "Socrate è mortale" è vera sse Soddisfa(Riferimento("Socrate"), "mortale") sse Soddisfa(Socrate, "mortale") sse Socrate è mortale

Regole per connettivi logici:

- "¬φ" è vera sse "φ" non è vera
- "φ ∧ ψ" è vera sse "φ" è vera e "ψ" è vera
- "φ ∨ ψ" è vera sse "φ" è vera o "ψ" è vera
- "φ → ψ" è vera sse "φ" non è vera o "ψ" è vera

Queste regole permettono di derivare le condizioni di verità di frasi complesse:

"Maria corre e Giovanni dorme" è vera sse "Maria corre" è vera e "Giovanni dorme" è vera sse Maria corre e Giovanni dorme

Regole per quantificatori:

Davidson seguì l'approccio di Tarski per i quantificatori. Per frasi della forma "Ogni F è G":

"Ogni F è G" è vera sse per ogni oggetto x, se x soddisfa "F" allora x soddisfa "G" sse $\forall x(F(x) \rightarrow G(x))$

Per frasi della forma "Qualche F è G" o "Un F è G":

"Un F è G" è vera sse esiste un oggetto x tale che x soddisfa "F" e x soddisfa "G" sse $\exists x(F(x) \wedge G(x))$

Davidson sviluppò particolarmente la sua teoria per trattare le frasi di azione con modificatori avverbiali. Consideriamo:

(10) Bruto pugnalò Cesare (11) Bruto pugnalò Cesare violentemente
(12) Bruto pugnalò Cesare violentemente nel Foro (13) Bruto pugnalò Cesare violentemente nel Foro alle idi di marzo

Una teoria semantica ingenua potrebbe trattare questi come predicati di arità crescente:

- Pugnalò(Bruto, Cesare)
- Pugnalò-violentemente(Bruto, Cesare)
- Pugnalò-violentemente-nel-Foro(Bruto, Cesare)
- Pugnalò-violentemente-nel-Foro-alle-idi-di-marzo(Bruto, Cesare)

Ma questo approccio fallisce nel catturare le inferenze logiche valide. Da (13) dovremmo poter inferire validamente (12), (11) e (10). Ma se questi sono predicati completamente diversi, non c'è connessione logica tra loro che giustifichi queste inferenze.

Davidson propose invece una quantificazione esistenziale su eventi. La forma logica di (10) è:

(10') $\exists$ e [Pugnalamento(e) $\wedge$ Agente(e, Bruto) $\wedge$ Paziente(e, Cesare)]

"Esiste un evento e tale che e è un pugnalamento, l'agente di e è Bruto, e il paziente (oggetto) di e è Cesare."

La forma logica di (13) diventa:

(13') $\exists$ e [Pugnalamento(e) $\wedge$ Agente(e, Bruto) $\wedge$ Paziente(e, Cesare) $\wedge$ Violento(e) $\wedge$ Luogo(e, Foro) $\wedge$ Tempo(e, idi-di-marzo)]

Ora le inferenze sono logicamente valide: (13') implica logicamente (10') perché ogni evento che soddisfa tutti i congiunti di (13') soddisfa anche i congiunti di (10'). I modificatori avverbiali sono semplicemente predicati addizionali sull'evento.

Questa analisi davidsoniana degli eventi è diventata standard nella semantica formale ed è stata estesa e raffinata da numerosi semanticisti, incluso Terence Parsons che ha sviluppato una teoria neo-davidsoniana ancora più fine-grained.

L'approccio davidsoniano ha il notevole vantaggio di fornire un framework formalmente rigoroso e sistematicamente compositivo per la semantica. La teoria è esplicita, testabile, e fa predizioni precise sulle condizioni di verità delle frasi. Tuttavia, affronta diverse difficoltà serie e profonde quando tenta di rendere conto della comprensione linguistica ordinaria nelle lingue naturali.

Problemi Fondamentali della Semantica Vero-Condizionale

Il Problema degli Enunciati Non-Dichiarativi

La semantica vero-condizionale standard, nella sua formulazione classica, si applica esclusivamente a frasi dichiarative che possono essere valutate come vere o false—enunciati che hanno lo scopo illocutivo di asserire che il mondo è in un certo modo.

Ma le lingue naturali contengono una ricca varietà di altri tipi di enunciati che servono funzioni comunicative completamente diverse:

Interrogativi (domande):

- "Che ore sono?"
- "Dove abita Maria?"
- "Chi ha scritto 'Guerra e Pace'?"
- "Perché il cielo è azzurro?"
- "Piove?"

Imperativi (comandi, richieste):

- "Chiudi la porta!"
- "Non correre nei corridoi!"
- "Per favore passami il sale"
- "Sia x tale che x > 5" (imperativo matematico)

Esclamativi:

- "Che bella giornata!"
- "Come sei cresciuto!"
- "Quanto è stupido!"

Ottativi (espressioni di desiderio):

- "Magari vincessi la lotteria!"
- "Se solo fossi più giovane!"

Nessuna di queste categorie di enunciati ha condizioni di verità nel senso ordinario e standard. Non ha alcun senso coerente dire che la domanda "Che ore sono?" è vera o falsa—le domande non sono il tipo di cose che possono essere vere o false. Eppure li comprendiamo perfettamente; sappiamo esattamente cosa significano questi enunciati e come usarli appropriatamente. Un parlante competente che ascolta "Chiudi la porta!" comprende immediatamente che gli viene richiesto di compiere una certa azione. Cosa significa comprendere una domanda, un comando o un'esclamazione se la comprensione non può consistere nella conoscenza delle condizioni di verità?

Alcuni teorici hanno proposto di generalizzare la nozione di condizioni di verità a quella più ampia di condizioni di soddisfacimento (satisfaction conditions) o condizioni di felicità (felicity conditions). Un imperativo come "Chiudi la porta!" sarebbe soddisfatto se e solo se l'ascoltatore chiude effettivamente la porta in un intervallo di tempo ragionevole dopo aver ricevuto il comando. Un interrogativo polare come "Piove?" sarebbe soddisfatto da una risposta che specifica correttamente se piove o no. Un interrogativo wh come "Che ore sono?" sarebbe soddisfatto da una risposta che specifica correttamente l'ora attuale.

Formalmente, potremmo tentare di definire:

Per un imperativo Imp(φ), dove φ è una formula che descrive un'azione: Soddisfatto(Imp(φ)) sse l'ascoltatore compie l'azione descritta da φ entro un tempo appropriato

Per un interrogativo polare Int-polar(φ): Soddisfatto(Int-polar(φ)) sse viene fornita una risposta che specifica correttamente se φ è vera o falsa

Per un interrogativo wh Int-wh(x, φ):
Soddisfatto(Int-wh(x, φ)) sse viene fornita una risposta che specifica un valore per x tale che φ(x) è vera

Ma questa manovra teorica, per quanto ingegnosa, non risolve il problema di fondo. Le condizioni di soddisfacimento per imperativi e interrogativi non sono semplicemente condizioni di verità con un altro nome. La relazione tra un imperativo e le sue condizioni di soddisfacimento è di natura profondamente diversa dalla relazione tra una frase dichiarativa e le sue condizioni di verità.

Una frase dichiarativa come "La neve è bianca" è vera se un certo stato di cose ottiene nel mondo (la neve essere bianca), indipendentemente da qualsiasi azione o risposta di un ascoltatore. Le condizioni di verità sono conditions-on-the-world, fatti circa come stanno le cose oggettivamente. Un imperativo come "Chiudi la porta!", al contrario, non descrive uno stato di cose esistente ma prescrive uno stato di cose che dovrebbe essere realizzato attraverso l'azione dell'ascoltatore. Le sue "condizioni di soddisfacimento" non sono fatti indipendenti ma dipendono essenzialmente dall'azione dell'ascoltatore in risposta al comando. Questo è un tipo completamente diverso di contenuto semantico.

Similmente, un interrogativo non descrive né prescrive ma richiede informazione. La sua "condizione di soddisfacimento" è che una certa lacuna epistemica venga colmata attraverso una risposta appropriata. Ancora una volta, questo è un tipo di contenuto semantico non riducibile a condizioni di verità.

Alcuni semanticisti, inclusi Jeroen Groenendijk e Martin Stokhof, hanno sviluppato teorie più sofisticate che trattano il significato delle domande come partizioni dello spazio logico. Un interrogativo polare come "Piove?" partizioni le possibilità in due celle: {mondi dove piove} e {mondi dove non piove}. Un interrogativo wh come "Chi corre?" partiziona le possibilità in base a quale individuo corre. Comprendere la domanda significa conoscere questa partizione; rispondere alla domanda significa specificare in quale cella della partizione si trova il mondo attuale.

Formalmente, per un set W di mondi possibili: $[\![\text{"Piove?"}]\!] = \{W_1, W_2\}$ dove $W_1 = \{w \in W \mid \text{piove in } w\}$ e $W_2 = \{w \in W \mid \text{non piove in } w\}$

$[\![\text{"Chi corre?"}]\!] = \{\{w \in W \mid \text{corre}(a) \text{ in } w\} \mid a \in D\}$, dove D è il dominio degli individui

Questa teoria delle domande come partizioni è elegante e matematicamente rigorosa, ma resta la questione se catturi veramente ciò che un parlante ordinario conosce quando comprende una domanda.

Il Problema del Contesto e dell'Indicialità Pervasiva

Un secondo problema profondo per la semantica vero-condizionale è che quasi tutte le espressioni linguistiche del linguaggio ordinario quotidiano contengono elementi il cui

riferimento o interpretazione dipende crucialmente dal contesto di enunciazione. Gli elementi indicali e dimostrativi più ovvi e classicamente riconosciuti includono:

Pronomi personali:

- io, tu, noi, voi
- me, te, lui, lei, loro

Dimostrativi:

- questo, quello, codesto
- qui, lì, là
- ora, allora

Avverbi temporali:

- ora, adesso, oggi, ieri, domani, dopodomani
- l'anno scorso, il mese prossimo

Tempi verbali:

- passato prossimo, passato remoto
- futuro semplice, futuro anteriore
- condizionale presente e passato

Consideriamo la frase semplice e apparentemente non problematica:

(14) Io ho fame

Le condizioni di verità di questa frase non possono essere specificate in modo assoluto, context-indipendente, perché dipendono strettamente da chi pronuncia la frase e quando la pronuncia. Se pronunciata da Maria alle ore 12:00 del 15 marzo 2024, la frase è vera se e solo se Maria ha fame alle 12:00 del 15 marzo 2024. Se pronunciata da Giovanni alle ore 18:00 del 20 aprile 2024, la stessa sequenza di parole "io ho fame" esprime condizioni di verità completamente diverse: è vera sse Giovanni ha fame alle 18:00 del 20 aprile 2024.

Possiamo tentare di accomodare questo fenomeno relativizzando esplicitamente la verità non alla frase tipo (type) astratta, ma alle sue occorrenze enunciative concrete (tokens) in contesti specifici:

Un'occorrenza enunciativa u della frase "Io ho fame" è vera sse il parlante di u ha fame al tempo di u.

Formalmente, se u è un'enunciazione (utterance) caratterizzata da:

- Parlante(u) = la persona che produce u
- Tempo(u) = il momento in cui u è prodotta
- Luogo(u) = la locazione dove u è prodotta

Allora: Vera(u, "Io ho fame") sse Ha-fame(Parlante(u), Tempo(u))

David Kaplan, nel suo lavoro seminale "Demonstratives" del 1977 (pubblicato nel 1989), formalizzò rigorosamente questa idea distinguendo tra due livelli semantici per espressioni indicali e dimostrative:

Carattere (Character): Una funzione da contesti a contenuti (intensioni). Il carattere è il significato linguistico convenzionale e stabile dell'espressione indicale.

Contenuto (Content): Una funzione da circostanze di valutazione (mondi possibili, tempi) a estensioni. Il contenuto è ciò che viene espresso da una particolare occorrenza dell'espressione in un particolare contesto.

Formalmente, un contesto c può essere modellato come una tupla: $c = \langle s, a, l, t, w \rangle$

dove:

- s = agente/parlante del contesto
- a = ascoltatore/destinatario
- l = luogo/locazione del contesto
- t = tempo del contesto
- w = mondo possibile del contesto

Per l'indicale "io", il carattere è la regola semantica: Carattere("io") = λc . Parlante(c)

Questa è una funzione che, dato un contesto c qualsiasi, restituisce come contenuto il parlante in quel contesto. Il contenuto risultante è a sua volta una funzione costante: Contenuto_c("io") = λ⟨w',t'⟩ . Parlante(c)

Questa è una funzione che, data qualsiasi circostanza di valutazione ⟨w',t'⟩, restituisce sempre lo stesso individuo: il parlante nel contesto c. Questo cattura il fatto che "io" è un designatore rigido—si riferisce allo stesso individuo in tutti i mondi possibili e tempi, una volta che il contesto è fissato.

Similmente: Carattere("tu") = λc . Ascoltatore(c)
Carattere("qui") = λc . Luogo(c) Carattere("ora") = λc . Tempo(c) Carattere("oggi") = λc . il-giorno-contenente(Tempo(c)) Carattere("ieri") = λc . il-giorno-precedente(il-giorno-contenente(Tempo(c)))

Per un dimostrativo come "questo" o "quello", Kaplan sosteneva che il carattere richiede anche un'intenzione dimostrativa o un pointing gesture per determinare il riferimento:

Carattere("questo") = λc . l'oggetto dimostrato dal parlante in c

Ma il problema dell'indicialità e della dipendenza dal contesto va molto, molto più in profondità di questi casi ovvi ed espliciti. Consideriamo:

(15) Ho fatto colazione

In condizioni normali di enunciazione, questa frase viene compresa come significante che il parlante ha fatto colazione nella mattinata dello stesso giorno dell'enunciazione—non ieri, non la settimana scorsa, ma stamattina. Ma questa restrizione temporale specifica non è

codificata esplicitamente in alcun elemento lessicale della forma linguistica. Il passato prossimo italiano indica semplicemente che l'evento di fare colazione precede il tempo di enunciazione, senza specificare quanto tempo prima. La restrizione temporale più specifica secondo cui l'evento debba essere avvenuto nella mattinata dello stesso giorno dell'enunciazione emerge non dalla semantica convenzionale codificata nella morfologia verbale del tempo passato prossimo, ma attraverso un processo di arricchimento pragmatico che opera al di sotto della soglia della consapevolezza cosciente del parlante, trasformando quello che sarebbe altrimenti un contenuto proposizionale radicalmente sottodeterminato in un messaggio comunicativamente adeguato e contestualmente appropriato.

Robyn Carston e altri teorici della pertinenza sostengono che questo fenomeno di arricchimento pragmatico non rappresenta un'aggiunta post-hoc opzionale al contenuto semantico già completamente determinato, ma costituisce invece una componente essenziale e costitutiva di ciò che il parlante comunica esplicitamente attraverso l'enunciazione. La distinzione tradizionale griceana tra "ciò che è detto" (what is said) e "ciò che è implicato conversazionalmente" (what is conversationally implicated), sebbene rimanga teoricamente preziosa come prima approssimazione, si rivela insufficientemente fine-grained per catturare la complessità empirica dell'interfaccia semantica-pragmatica. Carston introduce il concetto tecnico di "explicature" per designare quella rappresentazione proposizionale che risulta dall'arricchimento pragmatico della forma logica sottodeterminata generata dal sistema sintattico-semantico compositivo. L'explicature non è identica al contenuto semantico minimale codificato linguisticamente, né è un'implicatura nel senso griceano standard, ma occupa uno spazio teorico intermedio che richiede una riconfigurazione fondamentale della nostra concezione dell'architettura cognitiva della comprensione linguistica.

Per formalizzare questa distinzione tripartita con maggiore precisione, possiamo rappresentare il processo di interpretazione come una funzione composta che opera attraverso molteplici livelli rappresentazionali. Sia L la forma linguistica superficiale di un enunciato, ovvero la sequenza di espressioni fonologiche o ortografiche come effettivamente prodotte dal parlante. Il sistema sintattico-semantico compositivo opera su L applicando le regole della grammatica generativa e della semantica compositiva per derivare una rappresentazione logica sottostante che possiamo chiamare la forma logica minimale, denotata come LF(L). Questa forma logica minimale è una struttura proposizionale schematica che contiene variabili libere, sottodeterminazioni referenziali, ambiguità lessicali non risolte, e restrizioni di dominio non saturate.

Formalmente, possiamo concepire LF(L) come una formula del linguaggio della logica del primo ordine arricchita con operatori intensionali, indicali, e parametri contestuali non vincolati. Per esempio, per l'enunciato "Ho fatto colazione", la forma logica minimale potrebbe essere rappresentata come:

LF("Ho fatto colazione") = $\exists e \exists t$ [Fare-colazione(e) $\wedge$ Agente(e, io*) $\wedge$ Tempo(e, t) $\wedge$ $t < t^* \wedge R(t, t^*)$]

dove io* è un parametro indicale che deve essere saturato con il riferimento al parlante nel contesto di enunciazione, t* è il parametro temporale che corrisponde al tempo di enunciazione, e R è una relazione temporale sottodeterminata che specifica soltanto che il

tempo dell'evento precede il tempo di enunciazione senza determinare quanto recentemente l'evento si sia verificato. Il simbolo R potrebbe essere istanziato come "nella stessa giornata di", "nella stessa settimana di", "nello stesso anno di", o qualsiasi altra relazione temporale compatibile con il contenuto semantico minimale codificato dalla morfologia del passato prossimo.

Il processo di arricchimento pragmatico può essere modellato come una funzione Enrich che prende come argomenti la forma logica minimale LF(L), il contesto di enunciazione C (che include informazioni sul parlante, l'ascoltatore, il tempo, il luogo, le assunzioni di sfondo condivise, il common ground conversazionale, e qualsiasi altra informazione contestualmente saliente), e restituisce come output una proposizione completamente saturata e arricchita che Carston chiama l'explicature dell'enunciato:

Explicature(L, C) = Enrich(LF(L), C)

Nel caso del nostro esempio, il processo di arricchimento pragmatico saturerà il parametro indicale io* con il riferimento al parlante effettivo, risolverà la relazione temporale sottodeterminata R selezionando l'interpretazione temporalmente più ristretta compatibile con le aspettative di pertinenza, e produrrà una proposizione del tipo:

Explicature("Ho fatto colazione", C) = ∃e [Fare-colazione(e) ∧ Agente(e, Parlante(C)) ∧ Tempo(e) ∈ Mattina-di(Giorno(Tempo(C)))]

dove Mattina-di è una funzione che mappa ogni giorno all'intervallo temporale che costituisce la mattina di quel giorno secondo le convenzioni culturali rilevanti nel contesto C.

Questo processo di derivazione dell'explicature non è arbitrario o idiosincratico ma è governato sistematicamente dal Principio Comunicativo di Pertinenza, che costituisce il pilastro fondamentale della Teoria della Pertinenza sviluppata da Dan Sperber e Deirdre Wilson. Secondo questo principio, ogni atto di comunicazione ostensiva comunica automaticamente e inevitabilmente una presunzione della propria pertinenza ottimale per il destinatario. Più precisamente, quando un parlante produce intenzionalmente e ostentatamente un enunciato con l'intenzione manifesta di comunicare qualcosa a un ascoltatore, questa stessa produzione ostensiva garantisce implicitamente che l'enunciato è sufficientemente pertinente da meritare lo sforzo di elaborazione cognitiva che l'ascoltatore dovrà investire per interpretarlo, e che è il più pertinente che il parlante era in grado e disposto a produrre date le sue competenze comunicative e i suoi obiettivi informativi.

La nozione di pertinenza, a sua volta, è definita formalmente in termini di due dimensioni ortogonali che interagiscono in modo inversamente proporzionale. Un input informativo I è pertinente per un individuo S in un contesto cognitivo K nella misura in cui la sua elaborazione produce effetti cognitivi positivi in K e richiede uno sforzo di elaborazione minimale. Gli effetti cognitivi positivi includono qualsiasi modifica del sistema di credenze di S che costituisca un miglioramento cognitivo genuino, ovvero che incrementi la conoscenza di S circa il mondo. Tali effetti includono l'aggiunta di nuove assunzioni che si combinano con assunzioni già presenti in K per generare implicazioni contestuali che non sarebbero derivabili senza l'input I, il rafforzamento di assunzioni esistenti attraverso evidenza addizionale, o l'eliminazione di contraddizioni mediante la revisione di assunzioni precedentemente sostenute ma ora riconosciute come false o dubbie.

Formalmente, possiamo definire gli effetti cognitivi positivi di un input I in un contesto K come:

Effetti-Cognitivi-Positivi(I, K) = |{p | p ∈ Implicazioni-Contestuali(K ∪ {I}) ∧ p ∉ Implicazioni-Contestuali(K)}| + Σ[p ∈ K] Rafforzamento(p, I) - Σ[p ∈ K] Indebolimento(p, I)

dove Implicazioni-Contestuali(X) denota l'insieme di tutte le proposizioni derivabili dall'insieme di assunzioni X attraverso processi deduttivi e inferenziali, Rafforzamento(p, I) misura quanto l'input I incrementa la forza epistemica dell'assunzione p nel sistema di credenze, e Indebolimento(p, I) misura quanto I riduce la forza di p, potenzialmente portando alla sua eliminazione se la contraddizione è sufficientemente forte.

Lo sforzo di elaborazione, dall'altra parte, può essere concepito come una funzione delle risorse cognitive attenzionali, di working memory, e di processing time richieste per derivare gli effetti cognitivi dall'input. Maggiore è la complessità computazionale dell'integrazione dell'input con il contesto esistente, maggiore sarà lo sforzo richiesto. Formalmente:

Sforzo-di-Elaborazione(I, K) = α · Tempo-di-Processing(I, K) + β · Carico-Working-Memory(I, K) + γ · Risorse-Attenzionali(I, K)

dove α, β, γ sono parametri di ponderazione che riflettono i costi relativi delle diverse dimensioni dello sforzo cognitivo nell'architettura cognitiva umana.

La pertinenza può quindi essere definita come il rapporto tra effetti cognitivi e sforzo di elaborazione:

Pertinenza(I, K) = Effetti-Cognitivi-Positivi(I, K) / Sforzo-di-Elaborazione(I, K)

Questa definizione cattura l'intuizione che la pertinenza aumenta proporzionalmente agli effetti cognitivi e inversamente proporzionalmente allo sforzo, raggiungendo il massimo quando gli effetti sono massimali e lo sforzo è minimale.

Il Principio Comunicativo di Pertinenza può ora essere formulato con maggiore precisione nel modo seguente. Ogni atto di comunicazione ostensiva crea nell'ascoltatore l'aspettativa che esista almeno un'interpretazione I dell'enunciato tale che I è sufficientemente pertinente da giustificare lo sforzo di elaborazione investito dall'ascoltatore, e tale che I è la più pertinente interpretazione compatibile con le capacità e le preferenze del comunicatore. Questa aspettativa di pertinenza ottimale guida il processo di comprensione in modo altamente efficiente, permettendo all'ascoltatore di seguire un percorso euristico di minimo sforzo nell'esplorazione dello spazio delle interpretazioni possibili, testando ipotesi interpretative in ordine di accessibilità decrescente e arrestando la ricerca non appena viene identificata un'interpretazione che soddisfa le aspettative di pertinenza.

Formalmente, l'ascoltatore cerca l'interpretazione I* che soddisfa:

I* = argmax[I ∈ Interpretazioni-Accessibili(L, C)] Pertinenza(I, K)

soggetto al vincolo che Pertinenza(I*, K) ≥ θ per qualche soglia di accettabilità θ determinata dalle aspettative di pertinenza create dall'atto comunicativo ostensivo del parlante.

Questo framework teorico fornisce una spiegazione sistematica e predittiva del fenomeno dell'arricchimento pragmatico. Nel nostro esempio "Ho fatto colazione", molteplici interpretazioni temporali sono in principio accessibili, corrispondenti a diverse istanziazioni della relazione temporale sottodeterminata R. L'ascoltatore potrebbe in linea di principio interpretare l'enunciato come significante che il parlante ha fatto colazione in qualche momento del passato remoto, magari la settimana scorsa, o il mese scorso, o l'anno scorso. Tuttavia, tali interpretazioni temporalmente non ristrette, sebbene semanticamente compatibili con il contenuto codificato linguisticamente, sarebbero comunicativamente aberranti perché non genererebbero effetti cognitivi positivi sufficienti a giustificare lo sforzo di elaborazione. Se il parlante avesse voluto comunicare semplicemente che ha fatto colazione in qualche momento indeterminato del passato, avrebbe violato la massima griceana di Quantità fornendo informazione insufficiente, oppure avrebbe potuto utilizzare espressioni temporali esplicite per disambiguare l'intervallo temporale rilevante.

L'interpretazione arricchita secondo cui il parlante ha fatto colazione nella mattinata dello stesso giorno, al contrario, è altamente pertinente perché fornisce informazione temporalmente specifica che è plausibilmente rilevante per gli scopi conversazionali in corso, richiede uno sforzo di elaborazione minimale essendo l'interpretazione temporale più saliente e accessibile data la semantica convenzionale del passato prossimo in italiano (che tipicamente denota eventi recenti con continuata rilevanza per il presente), e rispetta la presunzione che il parlante sta cooperando comunicativamente fornendo esattamente la quantità di informazione necessaria per i suoi scopi comunicativi.

Questa analisi rivela che il confine tra semantica e pragmatica non può essere tracciato in modo netto e categorico alla superficie grammaticale del linguaggio. Le espressioni linguistiche codificano schemi proposizionali sottodeterminati che fungono da istruzioni procedurali per la costruzione di rappresentazioni proposizionali complete attraverso processi inferenziali pragmatici guidati da principi di ottimizzazione della pertinenza. Il contenuto comunicato esplicitamente da un enunciato, ovvero la sua explicature, è il prodotto dell'interazione dinamica tra informazione codificata linguisticamente e processi di arricchimento contestuale, e non può essere identificato semplicemente con il contenuto semantico minimale decodificato compositivamente dalla forma linguistica.

Questa prospettiva teorica ha implicazioni profonde per la comprensione della natura della competenza linguistica e della sua implementazione cognitiva. Se la comprensione linguistica richiede sistematicamente processi di arricchimento pragmatico che vanno ben oltre la decodifica semantica compositiva, allora il sistema cognitivo responsabile della comprensione non può consistere semplicemente in un modulo sintattico-semantico incapsulato che mappa forme linguistiche su rappresentazioni semantiche attraverso regole compositive deterministiche. Piuttosto, la comprensione deve coinvolgere l'integrazione stretta e immediata di informazione linguistica con conoscenza enciclopedica, aspettative contestuali, teoria della mente, e capacità inferenziali generali.

Le evidenze psicolinguistiche sperimentali supportano questa concezione integrata dell'architettura della comprensione. Gli studi di potenziali evento-correlati che misurano l'attività elettrica cerebrale durante l'elaborazione di enunciati pragmaticamente arricchiti rivelano che gli effetti dell'arricchimento pragmatico emergono con latenze temporali estremamente rapide, dell'ordine di trecento-quattrocento millisecondi dalla presentazione

dello stimolo critico, temporalmente sovrapposte agli effetti dell'elaborazione semantica lessicale riflessi nella componente N400. Questo suggerisce che i processi pragmatici non operano in una fase post-semantica tardiva dopo che la rappresentazione semantica compositiva è stata completamente costruita, ma piuttosto interagiscono dinamicamente e in tempo reale con i processi semantici durante il corso incrementale della comprensione.

Gli studi di eye-tracking durante la lettura forniscono evidenza convergente per questa integrazione temporalmente rapida di informazione semantica e pragmatica. Quando i lettori incontrano espressioni referenzialmente ambigue o temporalmente sottodeterminate, i pattern di movimenti oculari rivelano che i processi di risoluzione pragmatica iniziano immediatamente, entro poche centinaia di millisecondi dalla codifica percettiva della parola critica, senza attendere il completamento dell'analisi sintattica dell'intera clausola o frase. I tempi di fissazione su regioni testuali successive sono modulati dalle difficoltà di integrazione pragmatica in modo comparabile a come sono modulati dalle difficoltà di integrazione sintattica o semantica, suggerendo che tutti questi livelli di elaborazione contribuiscono in parallelo alla costruzione incrementale della rappresentazione interpretativa.

Questa evidenza empirica pone sfide sostanziali per le teorie modulari classiche dell'architettura cognitiva del linguaggio, secondo cui l'elaborazione linguistica procederebbe attraverso una sequenza di stadi discreti e informaticamente incapsulati, con il modulo sintattico che opera prima e indipendentemente dal modulo semantico, e con i processi pragmatici che intervengono solo tardivamente dopo che le rappresentazioni sintattiche e semantiche sono state completate. I dati suggeriscono invece un'architettura più interattiva in cui informazione di diversi livelli fluisce bidirezionalmente attraverso i componenti del sistema, con processi top-down guidati dalle aspettative contestuali e pragmatiche che influenzano concorrentemente l'elaborazione bottom-up guidata dall'input linguistico.

La modellazione computazionale di questa architettura interattiva rimane una sfida teorica e tecnica formidabile. I modelli connessionisti che implementano reti neurali artificiali con connessioni bidirezionali tra livelli rappresentazionali possono in principio catturare alcuni aspetti dell'interazione dinamica tra elaborazione semantica e pragmatica, ma spesso mancano della trasparenza esplicativa e della precisione predittiva quantitativa che caratterizzano i modelli simbolici basati su regole. I modelli probabilistici bayesiani che concepiscono la comprensione come inferenza probabilistica sotto incertezza offrono un framework formalmente elegante per integrare molteplici fonti di informazione pesate in base alle loro affidabilità relative, ma affrontano difficoltà computazionali nel trattare gli spazi di ipotesi combinatorialmente esplosivi che emergono quando si considera la gamma completa di interpretazioni possibili per enunciati linguisticamente complessi.

Una direzione promettente per la modellazione computazionale dell'interfaccia semantica-pragmatica è l'approccio dei modelli generativi probabilistici gerarchici, che rappresentano la comprensione come un processo di inferenza bayesiana in cui l'ascoltatore mantiene distribuzioni di probabilità su interpretazioni possibili a molteplici livelli di astrazione, aggiornando queste distribuzioni incrementalmente man mano che nuova informazione linguistica diventa disponibile, guidato da un modello generativo che specifica come le interpretazioni a livelli superiori generano probabilisticamente osservazioni linguistiche a livelli inferiori.

Modelli Bayesiani della Comprensione e Integrazione Probabilistica

Formalmente, un modello generativo bayesiano per la comprensione linguistica può essere specificato come una tripla ordinata $M = \langle H, P, L \rangle$, dove H è lo spazio delle ipotesi interpretative (che includono interpretazioni sintattiche, semantiche, e pragmatiche a diversi livelli di granularità), P è una distribuzione di probabilità prior su H che riflette le aspettative dell'ascoltatore prima di processare l'enunciato corrente (basate su conoscenza linguistica, conoscenza del mondo, contesto conversazionale, e assunzioni circa le intenzioni del parlante), e L è una funzione di verosimiglianza che specifica, per ogni ipotesi interpretativa $h \in H$, la probabilità che l'enunciato linguistico osservato e sarebbe prodotto se l'interpretazione corretta fosse h.

Data un'osservazione linguistica e, il compito dell'ascoltatore è inferire la distribuzione di probabilità posterior su interpretazioni applicando il teorema di Bayes:

$$P(h \mid e) = [P(e \mid h) \cdot P(h)] / P(e)$$

dove $P(h \mid e)$ è la probabilità posterior dell'interpretazione h dato l'enunciato osservato e, $P(e \mid h)$ è la verosimiglianza dell'enunciato e sotto l'interpretazione h, $P(h)$ è la probabilità prior di h, e $P(e) = \Sigma[h' \in H] P(e \mid h') \cdot P(h')$ è la probabilità marginale dell'enunciato, che funge da costante di normalizzazione.

L'interpretazione ottimale secondo il criterio della massima probabilità a posteriori (MAP, maximum a posteriori) è:

$$h_{MAP} = \text{argmax}[h \in H] P(h \mid e) = \text{argmax}[h \in H] P(e \mid h) \cdot P(h)$$

dove la seconda uguaglianza sfrutta il fatto che $P(e)$ è costante rispetto a h e quindi può essere ignorata nell'ottimizzazione.

Questo framework formale fornisce una base principiata per integrare informazione di diversi tipi nel processo di comprensione. La distribuzione prior $P(h)$ può incorporare bias derivanti da frequenze lessicali, preferenze strutturali sintattiche, aspettative semantiche basate su conoscenza del mondo, presunzioni pragmatiche di cooperatività comunicativa, e qualsiasi altra fonte di informazione che l'ascoltatore porta al compito interpretativo. La funzione di verosimiglianza $P(e \mid h)$, dal canto suo, modella quanto bene l'enunciato osservato "si adatta" a ciascuna interpretazione ipotizzata, catturando il fatto che alcune interpretazioni renderebbero la produzione dell'enunciato osservato altamente probabile mentre altre la renderebbero improbabile o impossibile.

Nel contesto specifico dell'arricchimento pragmatico temporale che abbiamo discusso, possiamo modellare il processo come segue. Lo spazio delle ipotesi H include tutte le possibili specificazioni temporali dell'evento di fare colazione, parametrizzate da un intervallo temporale t. La distribuzione prior $P(t)$ riflette le aspettative dell'ascoltatore circa quando il parlante potrebbe aver fatto colazione, basate su conoscenza delle convenzioni temporali del pasto, schemi temporali tipici delle attività quotidiane, e qualsiasi informazione contestuale specifica disponibile.

La funzione di verosimiglianza P("Ho fatto colazione" | t) cattura quanto sarebbe naturale o appropriato per il parlante usare il passato prossimo italiano per riferirsi a un evento di fare colazione occorso all'intervallo temporale t. Questa verosimiglianza sarà alta per intervalli temporali recenti, particolarmente la mattina dello stesso giorno, e decrescerà progressivamente per intervalli più remoti nel passato, riflettendo la semantica convenzionale del passato prossimo che tipicamente denota eventi recenti con continuata rilevanza presente.

Combinando il prior con la verosimiglianza attraverso la regola di Bayes, otteniamo una distribuzione posterior P(t | "Ho fatto colazione") che sarà fortemente concentrata sull'intervallo temporale corrispondente alla mattina dello stesso giorno, a meno che il contesto non fornisca evidenza specifica che sposta le aspettative verso interpretazioni alternative. Questo spiega matematicamente perché l'interpretazione arricchita emerge così rapidamente e automaticamente: è semplicemente la modalità della distribuzione posterior, ovvero l'interpretazione con la probabilità a posteriori massima dato l'enunciato osservato e il contesto.

Questo approccio bayesiano offre anche una spiegazione naturale per la flessibilità contestuale dell'arricchimento pragmatico. Se il contesto conversazionale rendesse saliente un intervallo temporale diverso, per esempio se la conversazione riguardasse eventi della settimana precedente, allora la distribuzione prior P(t) si sposterebbe per assegnare probabilità più alta a intervalli temporali corrispondenti alla settimana scorsa, e conseguentemente anche la distribuzione posterior si sposterebbe, potenzialmente producendo un'interpretazione diversa dell'enunciato "Ho fatto colazione" che lo ancora temporalmente alla settimana precedente piuttosto che alla mattina corrente.

Estendendo questa analisi bayesiana oltre il caso specifico dell'arricchimento temporale, possiamo concepire l'intero processo di comprensione linguistica come inferenza probabilistica gerarchica attraverso molteplici livelli rappresentazionali. Al livello più basso, l'ascoltatore deve inferire le parole intese dal parlante dato il segnale acustico o visivo osservato, un processo che richiede di risolvere ambiguità percettive, compensare per variabilità nella pronuncia, e segmentare il flusso continuo del segnale in unità lessicali discrete. Questo può essere modellato come inferenza bayesiana in cui le parole w costituiscono lo spazio delle ipotesi, il segnale percettivo s è l'osservazione, e la distribuzione posterior P(w | s) è computata combinando un modello acustico o ortografico P(s | w) che specifica quanto il segnale osservato è compatibile con ciascuna parola, con un modello di linguaggio P(w) che cattura le probabilità prior delle parole basate su frequenze lessicali e vincoli grammaticali.

Al livello successivo nella gerarchia, l'ascoltatore deve inferire la struttura sintattica dell'enunciato date le parole riconosciute. Questo richiede di risolvere ambiguità strutturali, assegnare ruoli grammaticali ai costituenti, e costruire una rappresentazione gerarchica che rispetta i vincoli della grammatica della lingua. Questo può essere modellato come inferenza bayesiana su strutture sintattiche T dato una sequenza di parole w, con la distribuzione posterior P(T | w) computata da un modello generativo probabilistico che specifica come le strutture sintattiche generano sequenze di parole attraverso regole di produzione probabilizzate, esattamente come nelle Grammatiche Probabilistiche Libere dal Contesto (PCFGs) che abbiamo discusso precedentemente.

Al livello semantico, l'ascoltatore deve inferire il significato proposizionale dell'enunciato data la struttura sintattica. Questo coinvolge risoluzione di ambiguità lessicali (disambiguazione dei sensi delle parole polisemiche), determinazione del riferimento di espressioni referenziali (pronomi, nomi propri, descrizioni definite), risoluzione dello scope dei quantificatori e altri operatori, e composizione dei significati delle parti secondo regole semantiche compositive. Anche questo può essere formulato come inferenza bayesiana in cui l'ascoltatore mantiene una distribuzione di probabilità su possibili rappresentazioni semantiche S data la struttura sintattica T, con la distribuzione P(S | T) determinata da un modello che specifica come le strutture sintattiche mappano su rappresentazioni semantiche attraverso regole di interpretazione probabilizzate.

Finalmente, al livello pragmatico, l'ascoltatore deve inferire l'intenzione comunicativa del parlante data la rappresentazione semantica dell'enunciato e il contesto conversazionale. Questo richiede arricchimento pragmatico per saturare sottodeterminazioni, derivazione di implicature conversazionali, riconoscimento di atti linguistici indiretti, e integrazione dell'informazione comunicata con il common ground conversazionale. Questo può essere modellato come inferenza bayesiana su intenzioni comunicative I data una rappresentazione semantica S e un contesto C, con la distribuzione P(I | S, C) computata attraverso un modello probabilistico che cattura come parlanti cooperativi razionali selezionano forme linguistiche per comunicare intenzioni date le loro credenze circa le credenze e le capacità inferenziali dell'ascoltatore.

La potenza esplicativa di questo framework bayesiano gerarchico risiede nella sua capacità di integrare naturalmente informazione di molteplici fonti e livelli, propagando incertezza attraverso i livelli della gerarchia in modo matematicamente principiato. L'incertezza irrisolta a livelli inferiori (per esempio, ambiguità percettiva nel riconoscimento di parole) si propaga verso l'alto influenzando le distribuzioni di probabilità a livelli superiori, mentre le aspettative e i vincoli a livelli superiori (per esempio, plausibilità semantica o aspettative pragmatiche) si propagano verso il basso attraverso le funzioni di verosimiglianza condizionata, influenzando le inferenze a livelli inferiori. Questo meccanismo di propagazione bidirezionale di informazione cattura naturalmente l'interattività osservata empiricamente tra diversi livelli di elaborazione linguistica.

La Semantica Formale dei Quantificatori e il Problema della Composizionalità Generalizzata

Ritornando ora dal dominio della pragmatica e dell'architettura cognitiva della comprensione alla semantica formale, è essenziale approfondire il trattamento compositivo dei quantificatori, che costituisce uno dei risultati tecnici più importanti e teoricamente illuminanti della semantica formale moderna. I quantificatori come "ogni", "qualche", "nessun", "la maggior parte", "pochi", e "molti" sollevano sfide profonde per la teoria della composizionalità perché non denotano entità semplici del dominio come fanno i nomi propri, né proprietà semplici come fanno i predicati monadici, ma piuttosto relazioni complesse di ordine superiore tra insiemi.

La semantica montagoviana che abbiamo delineato precedentemente tratta i sintagmi nominali quantificati come espressioni di tipo ⟨⟨e,t⟩,t⟩, ovvero come funzioni che prendono

proprietà (tipo ⟨e,t⟩) come argomenti e restituiscono valori di verità. Questa analisi può essere generalizzata e approfondita introducendo la teoria dei quantificatori generalizzati, sviluppata formalmente da Jon Barwise e Robin Cooper negli anni Ottanta come estensione rigorosa della logica del primo ordine standard per accomodare la ricchezza e complessità dei quantificatori del linguaggio naturale.

Un quantificatore generalizzato Q di tipo ⟨1⟩ è una funzione che mappa insiemi (l'estensione di predicati monadici) a valori di verità. Formalmente, dato un dominio D di entità, un quantificatore Q è un sottoinsieme della powerset di D:

$Q \subseteq P(D)$

dove P(D) denota l'insieme di tutti i sottoinsiemi di D. Il quantificatore Q è vero di un insieme $A \subseteq D$ esattamente quando $A \in Q$.

Per esempio, il quantificatore "tutto" può essere definito come:

$TUTTO = \{A \subseteq D \mid A = D\}$

ovvero l'insieme che contiene soltanto il dominio intero D. Il quantificatore "qualcosa" può essere definito come:

$QUALCOSA = \{A \subseteq D \mid A \neq \varnothing\}$

ovvero l'insieme di tutti i sottoinsiemi non-vuoti di D. Il quantificatore "niente" è il complemento di QUALCOSA:

$NIENTE = \{A \subseteq D \mid A = \varnothing\}$

ovvero l'insieme che contiene soltanto l'insieme vuoto.

Tuttavia, nella maggior parte dei casi nel linguaggio naturale, i quantificatori non operano su un singolo insieme ma piuttosto esprimono relazioni tra due insiemi, tipicamente chiamati il dominio di restrizione (restriction) e lo scope nucleare (nuclear scope). Per esempio, nella frase "Ogni studente corre", il quantificatore "ogni" mette in relazione l'insieme degli studenti (restrizione) con l'insieme dei corridori (scope nucleare), asserendo che il primo insieme è un sottoinsieme del secondo.

Per catturare questo fenomeno, definiamo quantificatori generalizzati di tipo ⟨1,1⟩ come relazioni binarie tra insiemi:

$Q \subseteq P(D) \times P(D)$

ovvero Q è un insieme di coppie ordinate di sottoinsiemi di D. Il quantificatore Q è vero della coppia ⟨A, B⟩ esattamente quando $\langle A, B\rangle \in Q$.

I quantificatori standard del linguaggio naturale possono ora essere definiti precisamente:

$OGNI \subseteq P(D) \times P(D)$ è definito da: $\langle A, B\rangle \in OGNI$ sse $A \subseteq B$

ovvero "Ogni A è B" è vero esattamente quando ogni elemento di A è anche elemento di B.

QUALCHE $\subseteq$ P(D) × P(D) è definito da: $\langle A, B\rangle \in$ QUALCHE sse $A \cap B \neq \varnothing$

ovvero "Qualche A è B" è vero esattamente quando l'intersezione di A e B è non-vuota.

NESSUN $\subseteq$ P(D) × P(D) è definito da: $\langle A, B\rangle \in$ NESSUN sse $A \cap B = \varnothing$

ovvero "Nessun A è B" è vero esattamente quando A e B sono disgiunti.

Questi sono quantificatori relativamente semplici che possono essere definiti in termini di operazioni insiemistiche elementari. Il potere della teoria dei quantificatori generalizzati emerge quando consideriamo quantificatori più complessi che non sono definibili nella logica del primo ordine standard. Per esempio, il quantificatore "la maggior parte":

LA-MAGGIOR-PARTE $\subseteq$ P(D) × P(D) è definito da: $\langle A, B\rangle \in$ LA-MAGGIOR-PARTE sse $|A \cap B| > |A \setminus B|$

ovvero "La maggior parte degli A sono B" è vero esattamente quando il numero di A che sono B supera il numero di A che non sono B. Equivalentemente, più della metà degli A sono B.

Similmente, possiamo definire quantificatori numerici proporzionali come "almeno n", "esattamente n", "al massimo n":

ALMENO-n $\subseteq$ P(D) × P(D): $\langle A, B\rangle \in$ ALMENO-n sse $|A \cap B| \geq n$

ESATTAMENTE-n $\subseteq$ P(D) × P(D): $\langle A, B\rangle \in$ ESATTAMENTE-n sse $|A \cap B| = n$

AL-MASSIMO-n $\subseteq$ P(D) × P(D): $\langle A, B\rangle \in$ AL-MASSIMO-n sse $|A \cap B| \leq n$

Un risultato formale importante è che mentre quantificatori come OGNI, QUALCHE, e NESSUN sono definibili nella logica del primo ordine (possono essere espressi usando $\forall$ e $\exists$ con i connettivi booleani), quantificatori come LA-MAGGIOR-PARTE e i quantificatori numerici proporzionali non sono definibili nella logica del primo ordine standard. Questo dimostra che il linguaggio naturale è semanticamente più espressivo della logica del primo ordine, e che una teoria semantica adeguata per il linguaggio naturale richiede risorse logiche che vanno oltre quelle della logica classica dei predicati.

La teoria dei quantificatori generalizzati permette anche di formulare rigorosamente proprietà universali dei quantificatori del linguaggio naturale, ovvero proprietà che tutti (o quasi tutti) i quantificatori nelle lingue umane soddisfano indipendentemente dalla particolare lingua. Queste proprietà costituiscono candidati per universali semantici che potrebbero riflettere vincoli cognitivi profondi sulla struttura della facoltà linguistica umana.

Una proprietà universale importante è la conservatività (conservativity). Un quantificatore Q è conservativo sse per tutti gli insiemi A, B:

$\langle A, B\rangle \in Q$ sse $\langle A, A \cap B\rangle \in Q$

Intuitivamente, un quantificatore è conservativo se il suo valore di verità dipende soltanto dagli elementi del dominio di restrizione A che soddisfano anche lo scope nucleare B,

ovvero dall'intersezione A ∩ B, e non dagli elementi di B che cadono fuori di A. Per esempio, OGNI è conservativo perché ⟨A, B⟩ ∈ OGNI (cioè A ⊆ B) se e solo se ⟨A, A ∩ B⟩ ∈ OGNI (cioè A ⊆ A ∩ B), e quest'ultima condizione è equivalente alla prima perché A ⊆ B implica A ∩ B = A.

La conservatività ha un significato linguistico profondo: essa implica che quando valutiamo la verità di "Ogni studente corre", dobbiamo soltanto considerare gli individui che sono studenti e verificare se questi corrono; non è rilevante quanti non-studenti corrono o non corrono. Questa proprietà sembra essere universale: tutti i determinanti quantificazionali nelle lingue naturali sono conservativi. Non esistono determinanti nella lingua naturale che violerebbero la conservatività, come il quantificatore ipotetico NON-CONSERVATIVO definito da:

⟨A, B⟩ ∈ NON-CONSERVATIVO sse |B \ A| > |A|

ovvero "NON-CONSERVATIVO A sono B" sarebbe vero sse ci sono più B che non sono A di quanti sono gli A in totale. Tale quantificatore, sebbene matematicamente definibile perfettamente, non corrisponde a nessun determinante in nessuna lingua naturale conosciuta.

Un'altra proprietà universale proposta è l'estensionalità (extensionality), che afferma che il valore di verità di un enunciato quantificato dipende soltanto dalle estensioni (gli insiemi) dei predicati coinvolti, e non dalle loro intensioni o dai modi particolari in cui quegli insiemi sono specificati. Formalmente, un quantificatore Q è estensionale sse:

Se A = A' e B = B', allora ⟨A, B⟩ ∈ Q sse ⟨A', B'⟩ ∈ Q

Questa proprietà garantisce che i quantificatori del linguaggio naturale si comportino in modo puramente set-theoretic, dipendendo soltanto da quali oggetti cadono sotto quali predicati e non da come quei predicati sono descritti o concettualizzati.

Una terza proprietà importante è la monotonia, che riguarda come il valore di verità di enunciati quantificati cambia quando espandiamo o restringiamo gli insiemi coinvolti. Un quantificatore Q è monotono crescente nella posizione destra (upward monotone in the right argument) sse:

Se ⟨A, B⟩ ∈ Q e B ⊆ B', allora ⟨A, B'⟩ ∈ Q

Intuitivamente, se "Q A sono B" è vero, e se B è un sottoinsieme di B', allora "Q A sono B'" deve anche essere vero. Il quantificatore QUALCHE è monotono crescente a destra: se "Qualche studente corre" è vero, e se tutti i corridori sono anche atleti, allora "Qualche studente è un atleta" deve anche essere vero.

Similmente, un quantificatore Q è monotono decrescente nella posizione destra (downward monotone) sse:

Se ⟨A, B'⟩ ∈ Q e B ⊆ B', allora ⟨A, B⟩ ∈ Q

Il quantificatore OGNI è monotono decrescente a destra nel senso duale: è monotono crescente nel suo argomento scope (se ogni studente corre e correre implica muoversi, allora ogni studente si muove), ma possiamo anche considerare la negazione.

Un fenomeno linguistico importante connesso alla monotonia è il licensing di elementi di polarità negativa (negative polarity items, NPIs). Gli NPIs sono espressioni come "mai", "alcuno" (nel senso negativo), "minimamente" che possono apparire soltanto in contesti con certe proprietà semantiche, tipicamente contesti monotoni decrescenti.

Elementi di Polarità Negativa e la Teoria Formale dei Contesti Monotoni Decrescenti

Gli elementi di polarità negativa (NPIs, negative polarity items) rappresentano un fenomeno linguistico affascinante che illustra in modo particolarmente trasparente l'interfaccia tra sintassi, semantica e processi inferenziali. In italiano, espressioni come "alcuno" nel suo uso negativo, "mai", "affatto", "il benché minimo", e locuzioni come "battere ciglio" o "spicccicare parola" sono NPIs che possono apparire soltanto in contesti sintattici e semantici con proprietà particolari. Consideriamo i seguenti contrasti di grammaticalità:

(16a) Giovanni non ha mai letto Proust. [grammaticale] (16b) *Giovanni ha mai letto Proust. [agrammaticale senza negazione]

(17a) Nessuno studente ha fatto alcun progresso. [grammaticale] (17b) *Ogni studente ha fatto alcun progresso. [agrammaticale]

(18a) Dubito che Maria abbia il benché minimo interesse per la fisica. [grammaticale] (18b) *Credo che Maria abbia il benché minimo interesse per la fisica. [agrammaticale]

Questi dati rivelano che gli NPIs richiedono un "licensor" che possieda certe proprietà semantiche specifiche. La negazione esplicita "non" in (16a) licenzia l'NPI "mai", mentre la sua assenza in (16b) rende la frase agrammaticale. Il quantificatore "nessuno" in (17a) licenzia "alcun", mentre il quantificatore "ogni" in (17b) non lo fa. Il verbo "dubitare" in (18a) licenzia l'NPI, mentre "credere" in (18b) non lo fa.

La generalizzazione semantica che unifica questi casi apparentemente eterogenei è stata formulata rigorosamente da William Ladusaw negli anni Ottanta: gli NPIs sono licenziati in contesti monotoni decrescenti (downward entailing contexts). Un contesto C[·] è monotono decrescente rispetto a una posizione se, per tutte le proposizioni p e q tali che p implica logicamente q, abbiamo:

C[q] implica logicamente C[p]

In altre parole, in un contesto monotono decrescente, sostituire un'espressione con una logicamente più forte (che implica la prima) produce un enunciato logicamente più debole (che è implicato dal primo).

Formalizziamo questa nozione con maggiore precisione. Dato un contesto linguistico C[·] con una posizione da riempire indicata dal simbolo ·, e date due espressioni α e β tali che

[[α]] ⊆ [[β]] (ovvero l'estensione di α è un sottoinsieme dell'estensione di β, il che significa che α implica β), il contesto C è monotono decrescente nella posizione · se e solo se:

[[C[β]]] ⊆ [[C[α]]]

ovvero se C[β] implica logicamente C[α].

Applichiamo questa definizione ai casi sopra. Consideriamo il contesto negativo "Giovanni non ha letto ·". Prendiamo α = "niente" e β = "Proust". Certamente [[niente]] ⊆ [[Proust]] (l'insieme vuoto è sottoinsieme di ogni insieme). Ora valutiamo:

C[Proust] = "Giovanni non ha letto Proust" C[niente] = "Giovanni non ha letto niente"

È evidente che "Giovanni non ha letto niente" implica "Giovanni non ha letto Proust": se Giovanni non ha letto assolutamente nulla, allora a fortiori non ha letto Proust. Quindi [[C[niente]]] ⊆ [[C[Proust]]], il che significa che [[C[β]]] ⊆ [[C[α]]], confermando che il contesto negativo è monotono decrescente.

Per contrasto, consideriamo il contesto affermativo "Giovanni ha letto ·". Di nuovo, prendiamo α = "niente" e β = "Proust". Valutiamo:

C[Proust] = "Giovanni ha letto Proust" C[niente] = "Giovanni ha letto niente"

Ma ora l'implicazione va nella direzione opposta: "Giovanni ha letto Proust" implica "Giovanni ha letto qualcosa" (negazione di "ha letto niente"), ma "Giovanni ha letto niente" (interpretato come "Giovanni non ha letto niente", ovvero "Giovanni ha letto qualcosa") non implica necessariamente "Giovanni ha letto Proust". Di fatto, la frase "Giovanni ha letto niente" è semanticamente anomala in questo contesto affermativo. Il contesto affermativo è quindi monotono crescente (upward entailing), non decrescente, e quindi non licenzia NPIs.

Consideriamo ora i quantificatori. Il quantificatore "nessuno" crea un contesto monotono decrescente nella sua restriction e nel suo scope nucleare. Per la restriction, consideriamo:

"Nessun [studente di logica] ha risolto il problema" "Nessun [studente] ha risolto il problema"

Dato che [[studente di logica]] ⊆ [[studente]], se nessun studente (in generale) ha risolto il problema, allora a fortiori nessuno studente di logica ha risolto il problema. L'implicazione va dall'espressione più generale a quella più specifica, confermando la monotonia decrescente.

Per il scope nucleare:

"Nessuno studente ha [risolto il problema difficile]" "Nessuno studente ha [risolto il problema]"

Dato che [[risolto il problema difficile]] ⊆ [[risolto il problema]], se nessuno studente ha risolto alcun problema, allora certamente nessuno studente ha risolto il problema difficile. Di nuovo, l'implicazione va dall'espressione più generale a quella più specifica.

Il quantificatore "ogni", per contrasto, è monotono decrescente nella sua restriction ma monotono crescente nel suo scope nucleare. Per la restriction:

"Ogni [studente di logica] ha risolto il problema" "Ogni [studente] ha risolto il problema"

Se ogni studente in generale ha risolto il problema, allora certamente ogni studente di logica ha risolto il problema. La monotonia è decrescente.

Ma per lo scope nucleare:

"Ogni studente ha [risolto il problema difficile]" "Ogni studente ha [risolto il problema]"

Se ogni studente ha risolto il problema difficile, allora certamente ogni studente ha risolto qualche problema. L'implicazione va dall'espressione più specifica a quella più generale, indicando monotonia crescente. Questo spiega perché "ogni" può licenziare NPIs nella sua restriction ma non nel suo scope nucleare:

(19a) Ogni studente che abbia mai studiato logica ha apprezzato il corso. [grammaticale - NPI nella restriction] (19b) *Ogni studente ha mai studiato logica. [agrammaticale - NPI nello scope]

La teoria dei contesti monotoni decrescenti come licensors di NPIs fornisce predizioni precise e testabili empiricamente circa la distribuzione di questi elementi nelle lingue naturali, e queste predizioni sono state ampiamente confermate attraverso studi tipologici cross-linguistici. Inoltre, questa teoria rivela una connessione profonda tra proprietà semantiche formali (monotonia nelle inferenze) e vincoli sintattici sulla distribuzione di elementi lessicali, suggerendo che la facoltà linguistica umana è sensibile a distinzioni logiche astratte che operano a livello della struttura inferenziale del linguaggio.

Le Implicature Scalari e la Grammaticalizzazione dell'Inferenza Pragmatica

Ritorniamo ora al fenomeno delle implicature scalari che abbiamo menzionato precedentemente nel contesto della teoria griceana della conversazione. Le implicature scalari rappresentano un caso particolarmente interessante di interfaccia semantica-pragmatica perché sollevano la questione fondamentale di dove tracciare il confine tra contenuto semantico codificato convenzionalmente nella lingua e contenuto pragmatico derivato inferenzialmente dal contesto comunicativo.

Ricordiamo l'esempio paradigmatico:

(20) Maria ha tre figli.

Nell'uso comunicativo ordinario, questo enunciato viene interpretato spontaneamente come comunicante che Maria ha esattamente tre figli, non più di tre. Tuttavia, il contenuto semantico letterale codificato dalle parole sembra essere soltanto che Maria ha almeno tre figli, compatibilmente con la possibilità che ne abbia quattro, cinque, o qualsiasi numero maggiore. La restrizione "esattamente tre" emerge come un'inferenza pragmatica derivata dal principio di cooperazione conversazionale.

Il ragionamento inferenziale può essere ricostruito formalmente come segue. I numerali "uno", "due", "tre", "quattro", eccetera formano una scala di Horn ordinata per forza informativa:

⟨uno, due, tre, quattro, cinque, ...⟩

dove ogni elemento della scala implica logicamente tutti gli elementi che lo precedono: se Maria ha quattro figli, allora certamente ha tre figli, due figli, e un figlio. Formalmente, se n e m sono numerali con n > m, allora:

[[Maria ha n figli]] ⊆ [[Maria ha m figli]]

ovvero l'insieme di mondi possibili in cui Maria ha n figli è un sottoinsieme dell'insieme di mondi in cui Maria ha m figli.

Quando un parlante cooperativo enuncia "Maria ha tre figli", l'ascoltatore ragiona come segue. Se il parlante fosse in posizione epistemica di asserire un'alternativa scalare più forte, ovvero "Maria ha quattro figli", "Maria ha cinque figli", o qualsiasi numerale superiore, allora il parlante avrebbe dovuto asserire quell'alternativa più forte in osservanza della massima griceana di Quantità (rendi il tuo contributo informativo quanto richiesto). Il fatto che il parlante ha invece asserito soltanto "Maria ha tre figli" implica conversazionalmente che il parlante non è in posizione di asserire alcuna delle alternative scalari più forti. Assumendo che il parlante sia competente e informato riguardo al dominio in questione, l'incapacità del parlante di asserire "Maria ha quattro figli" implica che "Maria ha quattro figli" è falso. Similmente per tutti i numerali superiori. Quindi, per esclusione, Maria ha esattamente tre figli.

Possiamo formalizzare questo ragionamento utilizzando un operatore di esaustificazione che modella l'inferenza pragmatica. Sia φ l'enunciato asserito dal parlante, e sia Alt(φ) l'insieme delle alternative scalari a φ, ovvero l'insieme di enunciati che il parlante avrebbe potuto asserire al posto di φ e che sono ordinati rispetto a φ sulla scala di Horn rilevante. Definiamo le alternative scalari più forti come:

Alt_+(φ) = {ψ ∈ Alt(φ) | [[ψ]] ⊂ [[φ]] e non esiste ψ' ∈ Alt(φ) tale che [[ψ]] ⊂ [[ψ']] ⊂ [[φ]]}

ovvero Alt_+(φ) contiene le alternative che sono logicamente più forti di φ (implicano φ ma non viceversa) e che sono massimali in questa proprietà (non esistono alternative intermedie nella scala).

L'operatore di esaustificazione Exh può essere definito come:

Exh(φ) = φ ∧ ∧[ψ ∈ Alt_+(φ)] ¬ψ

ovvero l'esaustificazione di φ è la congiunzione di φ con la negazione di tutte le sue alternative scalari più forti. Questo corrisponde precisamente all'implicatura scalare: asserire φ implica conversazionalmente che tutte le alternative più forti sono false.

Applicando questo al nostro esempio:

Alt_+("Maria ha tre figli") = {"Maria ha quattro figli", "Maria ha cinque figli", "Maria ha sei figli", ...}

Exh("Maria ha tre figli") = "Maria ha tre figli" ∧ ¬"Maria ha quattro figli" ∧ ¬"Maria ha cinque figli" ∧ ... = "Maria ha esattamente tre figli"

Questa derivazione formale cattura elegantemente l'intuizione che l'interpretazione "esattamente tre" emerge dalla combinazione del contenuto semantico letterale "almeno tre" con l'implicatura che tutti i numerali superiori sono falsi.

Tuttavia, una questione teorica profonda emerge quando esaminiamo il comportamento temporale e la resistenza alla cancellazione delle implicature scalari. Gli studi psicolinguistici sperimentali rivelano che le implicature scalari vengono computate con latenze temporali estremamente rapide, comparabili alle latenze dell'accesso semantico lessicale, e che sono difficili da sopprimere anche quando il contesto linguistico esplicito le contraddirebbe. Questo ha portato alcuni teorici, particolarmente nel framework della grammatica come Gennaro Chierchia, a proporre che le implicature scalari non sono derivate post-hoc attraverso ragionamento pragmatico generale ma sono invece grammaticalizzate, ovvero computate da un meccanismo sintattico-semantico specializzato che inserisce l'operatore di esaustificazione Exh come componente della forma logica dell'enunciato a livelli specifici della derivazione sintattica.

Secondo questa teoria neo-griceana grammaticalizzata, esisterebbe un operatore nullo Exh che può essere inserito opzionalmente in posizioni sintattiche specifiche, tipicamente sopra le proiezioni di determinanti quantificazionali o modificatori scalari, e che opera semanticamente esattamente come l'operatore di esaustificazione che abbiamo definito sopra, negando le alternative scalari più forti. La presenza o assenza di questo operatore nella struttura sintattica determinerebbe se l'implicatura scalare è generata o no.

Formalmente, la forma logica di "Maria ha tre figli" con l'operatore esaustivo inserito sarebbe:

[Exh [Maria ha tre figli]]

che si interpreta semanticamente come:

[[Exh [Maria ha tre figli]]] = [[Maria ha tre figli]] ∧ ∧[ψ ∈ Alt_+("Maria ha tre figli")] ¬[[ψ]]

producendo l'interpretazione "esattamente tre figli" direttamente a livello compositivo semantico piuttosto che attraverso inferenza pragmatica post-semantica.

La questione se le implicature scalari siano meglio analizzate come fenomeni genuinamente pragmatici derivati da principi generali di cooperazione conversazionale (l'approccio griceano classico) o come fenomeni grammaticalizzati computati da meccanismi sintattici-semantici specializzati (l'approccio neo-griceano grammaticalizzato) rimane oggetto di dibattito teorico intenso e di indagine empirica attiva nella semantica e psicolinguistica contemporanea. Le evidenze empiriche sono miste, con alcuni dati che supportano la flessibilità contestuale e la cancellabilità caratteristiche dell'inferenza pragmatica, e altri dati che rivelano automaticità e rapidità suggestive di computazione grammaticalizzata.

Una possibilità teorica conciliatoria è che esistano molteplici meccanismi per derivare interpretazioni esaustive, operanti a diversi livelli dell'architettura cognitiva della comprensione. Un meccanismo grammaticalizzato rapido e automatico, operante attraverso l'inserzione di operatori esaustivi nella forma logica, potrebbe generare interpretazioni scalari di default che vengono poi soggette a revisione attraverso processi pragmatici più lenti e più flessibili quando il contesto fornisce evidenza che l'interpretazione di default è inappropriata. Questa architettura duale combinerebbe l'efficienza computazionale della grammaticalizzazione con la flessibilità adattiva dell'inferenza pragmatica generale.

La Semantica delle Presupposizioni e il Problema della Proiezione

Un altro fenomeno semantico-pragmatico che rivela la complessità dell'interfaccia tra contenuto codificato linguisticamente e inferenza contestuale è il fenomeno delle presupposizioni. Le presupposizioni sono proposizioni che sono date per scontate, presupposte come parte del common ground conversazionale, piuttosto che asserite o messe in questione dall'enunciato. Per esempio:

(21) Il re di Francia è calvo.

Questo enunciato, come abbiamo discusso estensivamente nel contesto della teoria delle descrizioni di Russell, presuppone che esiste un re di Francia. Se questa presupposizione è falsa, ovvero se non esiste alcun re di Francia, l'enunciato non è semplicemente falso (come sosteneva Russell) ma piuttosto è in qualche modo difettoso o inadeguato, manifestando un "fallimento presupposizionale" (presupposition failure).

(22) Giovanni ha smesso di fumare.

Questo enunciato presuppone che Giovanni fumava precedentemente. Se questa presupposizione è falsa, l'enunciato è inappropriato anche se non è asserito esplicitamente che Giovanni fumava.

(23) È Maria che ha rotto il vaso.

Questo enunciato, una costruzione scisso-focale (cleft), presuppone che qualcuno ha rotto il vaso, e asserisce che questo qualcuno è Maria.

La caratteristica distintiva delle presupposizioni, che le distingue sia dal contenuto asserito sia dalle implicature conversazionali, è il loro comportamento sotto negazione e in contesti embedded. Consideriamo la negazione di (22):

(24) Giovanni non ha smesso di fumare.

Intuitivamente, (24) continua a presupporre che Giovanni fumava precedentemente, esattamente come (22). La negazione nega soltanto l'asserzione che Giovanni ha cessato questa attività, ma non cancella la presupposizione che l'attività esisteva. Questo contrasta marcatamente con il contenuto asserito, che è direttamente negato, e con le implicature conversazionali, che tipicamente scompaiono sotto negazione.

Similmente, le presupposizioni si proiettano (project) attraverso operatori modali e attitudinali:

(25) È possibile che il re di Francia sia calvo.

Anche questo enunciato modale presuppone l'esistenza del re di Francia.

(26) Giovanni crede che Maria abbia smesso di fumare.

Questo enunciato presuppone che Maria fumava, proiettando la presupposizione dall'embedded clause alla main clause.

Il problema della proiezione (projection problem) per le presupposizioni, formulato classicamente da Lauri Karttunen negli anni Settanta, è il problema di predire sistematicamente come le presupposizioni di clausole embedded si comportano quando queste clausole sono incorporate in costruzioni sintattiche più complesse. Alcune costruzioni linguistiche sono "holes" (buchi) rispetto alle presupposizioni, permettendo alle presupposizioni di proiettarsi intatte dalla embedded clause alla clausola complessiva. Altre costruzioni sono "plugs" (tappi), bloccando completamente la proiezione. Altre ancora sono "filters" (filtri), permettendo la proiezione condizionale su certe configurazioni strutturali e contestuali.

La negazione è un hole: le presupposizioni proiettano attraverso la negazione senza modifiche.

I verbi fattivi come "sapere", "rendersi conto", "rammaric arsi" sono holes:

(27) Giovanni sa che Maria ha smesso di fumare.

Presuppone sia che Maria fumava (presupposizione di "smettere") sia che Maria ha smesso di fumare (presupposizione fattiva di "sapere").

I verbi modali epistemici come "potere", "dovere" in lettura epistemica sono holes:

(28) Maria deve aver smesso di fumare.

Presuppone che Maria fumava.

Per contrasto, i verbi di dire (saying verbs) come "dire", "affermare", "sostenere" sono plugs, bloccando la proiezione:

(29) Giovanni dice che il re di Francia è calvo.

Questo enunciato non presuppone che esiste un re di Francia; asserisce soltanto che Giovanni fa questa affermazione, che potrebbe essere falsa o basata su presupposizioni erronee di Giovanni.

I connettivi condizionali manifestano un comportamento di filtering complesso. Consideriamo:

(30) Se Giovanni ha figli, allora tutti i suoi figli sono calvi.

L'antecedente "Giovanni ha figli" cancella la presupposizione esistenziale del conseguente "tutti i suoi figli sono calvi". Il condizionale nel suo insieme non presuppone che Giovanni abbia figli, perché questa informazione è fornita esplicitamente nell'antecedente.

Ma consideriamo:

(31) Se il re di Francia è calvo, allora il re di Francia è felice.

Qui la presupposizione esistenziale "esiste un re di Francia" viene proiettata dall'intero condizionale, anche se appare sia nell'antecedente sia nel conseguente.

Karttunen formulò una generalizzazione per i condizionali: una presupposizione p di un conseguente C si proietta da "Se A allora C" a meno che p sia implicata (entailed) dall'antecedente A o da A congiunto con il common ground. Formalmente:

Presupposizioni("Se A allora C") = Presupposizioni(A) ∪ {p ∈ Presupposizioni(C) | non A ⊨ p}

Questa predice correttamente che (30) non presuppone che Giovanni ha figli (perché l'antecedente lo implica esplicitamente), mentre (31) presuppone che esiste un re di Francia (perché l'antecedente "il re di Francia è calvo" non implica l'esistenza da solo secondo l'analisi russelliana, ma piuttosto la presuppone anch'esso).

Irene Heim sviluppò una teoria dinamica delle presupposizioni negli anni Ottanta che integra il trattamento delle presupposizioni con la semantica dinamica del discorso che abbiamo discusso precedentemente. Secondo la teoria di Heim, le presupposizioni sono vincoli di ammissibilità (admissibility constraints) sull'input context per l'interpretazione di un enunciato. Un enunciato con presupposizione p può essere enunciato felicemente (uttered felicitously) in un contesto C soltanto se C già entails p, ovvero se p è già parte del common ground rappresentato da C.

La Dinamica delle Presupposizioni nell'Aggiornamento del Common Ground

Formalmente, nella semantica di Heim, il significato di una frase non è una funzione statica da mondi possibili a valori di verità, ma piuttosto una funzione dinamica che mappa context sets di input a context sets di output. Un context set è un insieme di mondi possibili che rappresentano le possibilità epistemiche ancora aperte dato il common ground conversazionale corrente. Quando un enunciato viene asserito, il context set viene aggiornato eliminando tutti i mondi possibili in cui l'enunciato è falso, ristringendo così le possibilità epistemiche.

Sia C un context set, ovvero un insieme di mondi possibili. Una frase φ con presupposizione p è ammissibile (admissible) in C sse per ogni mondo w ∈ C, p è vera in w. Formalmente:

Admissible(φ, C) sse C ⊆ [[p]]

dove [[p]] è l'insieme di mondi in cui p è vera. Se questa condizione è soddisfatta, allora l'enunciazione di φ aggiorna il context set eliminando tutti i mondi in cui l'asserzione di φ è falsa:

C[φ] = C ∩ [[φ]]

dove [[φ]] è l'insieme di mondi in cui l'asserzione di φ è vera (distinta dalla sua presupposizione).

Se invece la condizione di ammissibilità non è soddisfatta, ovvero se esistono mondi nel context set corrente in cui la presupposizione p è falsa, allora l'enunciato φ non può essere interpretato felicemente e richiede un processo di accomodamento presupposizionale (presupposition accommodation). L'accomodamento consiste nell'aggiungere tacitamente la presupposizione al common ground, ristringendo il context set per eliminare i mondi in cui la presupposizione è falsa, prima di procedere con l'aggiornamento basato sull'asserzione:

Se non Admissible(φ, C), allora: C' = C ∩ [[p]] (accomodamento) C[φ] = C' ∩ [[φ]] (aggiornamento normale)

Questo meccanismo di accomodamento spiega come le presupposizioni possono essere introdotte nel discorso anche quando non sono già stabilite nel common ground. Quando un parlante enuncia "Il re di Francia è calvo" in un contesto dove l'esistenza del re di Francia non è stata previamente stabilita, l'ascoltatore può accomodare la presupposizione esistenziale, aggiungendola tacitamente al common ground se questa aggiunta è compatibile con le sue credenze e con lo scopo conversazionale.

Tuttavia, l'accomodamento non è sempre possibile o appropriato. Se la presupposizione contraddice esplicitamente informazione già stabilita nel common ground, o se contraddice la conoscenza del mondo dell'ascoltatore, allora l'accomodamento fallisce e l'enunciato viene percepito come infelicitous o pragmaticamente difettoso:

(32) #Il mio gatto unicorno è malato.

L'ascoltatore non può accomodare facilmente la presupposizione che il parlante possiede un gatto unicorno, dato che questa contraddice la conoscenza del mondo che i gatti unicorno non esistono.

La teoria dinamica delle presupposizioni di Heim fornisce anche una spiegazione elegante del problema della proiezione attraverso connettivi complessi. Consideriamo la congiunzione:

(33) Giovanni ha figli e tutti i suoi figli sono calvi.

Il primo congiunto "Giovanni ha figli" asserisce esattamente ciò che il secondo congiunto "tutti i suoi figli sono calvi" presuppone. Secondo la semantica dinamica, il context set viene aggiornato sequenzialmente da sinistra a destra. Dopo l'elaborazione del primo congiunto, il context set C viene ristretto a C_1 = C ∩ [[Giovanni ha figli]]. Ora, quando il secondo congiunto viene processato, la sua presupposizione "Giovanni ha figli" è già soddisfatta in C_1, quindi non c'è fallimento presupposizionale e non è richiesto accomodamento. La presupposizione è localmente soddisfatta e non proietta alla congiunzione complessiva.

Per contrasto, se invertiamo l'ordine dei congiunti:

(34) Tutti i figli di Giovanni sono calvi e Giovanni ha figli.

Ora il primo congiunto presuppone "Giovanni ha figli", ma questa presupposizione non è soddisfatta nel context set iniziale C (assumendo che non sia già stabilita). L'accomodamento globale aggiunge la presupposizione al context set iniziale, quindi l'intera congiunzione (34) presuppone che Giovanni ha figli, anche se questo è asserito esplicitamente dal secondo congiunto. Questo spiega l'asimmetria tra (33) e (34): (33) non presuppone che Giovanni ha figli (perché questa informazione è fornita dal primo congiunto prima che sia richiesta dal secondo), mentre (34) presuppone che Giovanni ha figli (perché questa informazione è richiesta dal primo congiunto prima di essere fornita dal secondo).

Questo fenomeno, chiamato "local accommodation" o "filtering incrementale", emerge naturalmente dalla semantica dinamica senza bisogno di stipulazioni ad hoc, ed è stato confermato empiricamente attraverso studi psicolinguistici che misurano le aspettative degli ascoltatori e le difficoltà di elaborazione in punti critici del discorso.

Anafora Discorsiva e la Teoria delle Rappresentazioni del Discorso

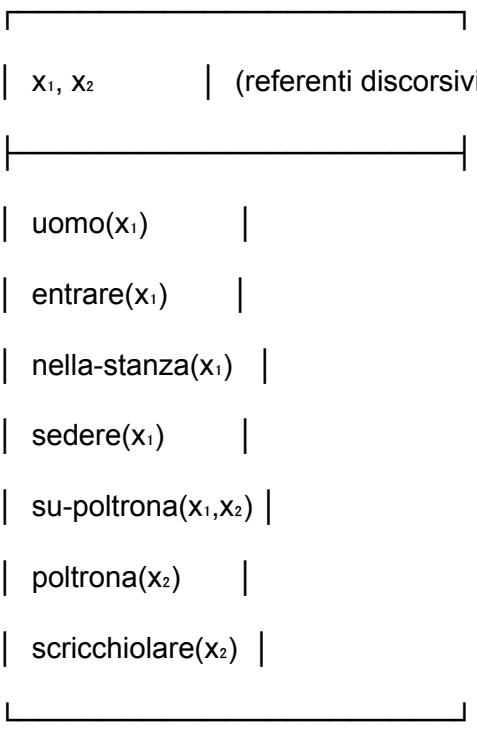
Il fenomeno dell'anafora discorsiva, ovvero l'uso di pronomi e altre espressioni anaforiche per riferirsi a entità introdotte precedentemente nel discorso, fornisce alcune delle evidenze più convincenti per l'adozione di una semantica dinamica che tratta il significato come potenziale di cambiamento del contesto piuttosto che come condizioni di verità statiche. Consideriamo il seguente mini-discorso:

(35) Un uomo$_1$ entrò nella stanza. Lui$_1$ si sedette sulla poltrona. La poltrona$_2$ scricchiolò.

L'interpretazione naturale di questo discorso coinvolge la costruzione di una rappresentazione mentale coerente in cui le espressioni anaforiche "lui$_1$" e "la poltrona$_2$" sono correttamente risolte ai loro antecedenti "un uomo$_1$" e "sulla poltrona". La Discourse Representation Theory (DRT) di Hans Kamp e la File Change Semantics di Irene Heim, entrambe sviluppate negli anni Ottanta, forniscono framework formali sistematici per modellare questo processo di costruzione incrementale di rappresentazioni discorsive.

Secondo la DRT, la comprensione di un discorso consiste nella costruzione progressiva di una Discourse Representation Structure (DRS), che è una struttura di dati complessa contenente due componenti principali: un insieme di referenti discorsivi (discourse referents), che sono variabili che rappresentano le entità introdotte nel discorso, e un insieme di condizioni (conditions) che specificano proprietà e relazioni tra questi referenti.

Una DRS può essere rappresentata formalmente come una coppia ordinata $K = \langle U_K, Con_K \rangle$, dove U_K è l'insieme dei referenti discorsivi e Con_K è l'insieme delle condizioni. Per il discorso (35), la DRS costruita incrementalmente avrebbe la seguente struttura finale:



La costruzione di questa DRS procede incrementalmente, frase per frase. L'elaborazione della prima frase "Un uomo entrò nella stanza" introduce un nuovo referente discorsivo x_1 nella componente U_K e aggiunge le condizioni uomo(x_1), entrare(x_1), nella-stanza(x_1) alla componente Con_K. Crucialmente, questo referente discorsivo x_1 rimane accessibile per riferimento anaforico nelle frasi successive.

L'elaborazione della seconda frase "Lui si sedette sulla poltrona" richiede la risoluzione del pronome "lui" a un referente discorsivo già presente nella DRS. Il sistema di risoluzione anaforica identifica x_1 come l'antecedente appropriato (essendo l'unico referente maschile singolare accessibile nel contesto), quindi la condizione sedere(x_1) viene aggiunta. Inoltre, la frase introduce implicitamente una nuova entità, la poltrona, che viene rappresentata mediante un nuovo referente discorsivo x_2 con le condizioni poltrona(x_2) e su-poltrona(x_1,x_2).

L'elaborazione della terza frase "La poltrona scricchiolò" risolve la descrizione definita "la poltrona" al referente x_2 già introdotto, e aggiunge la condizione scricchiolare(x_2).

Le regole di costruzione della DRT specificano formalmente, per ogni tipo di costruzione linguistica, come quella costruzione contribuisce all'aggiornamento della DRS. Per sintagmi nominali quantificati esistenzialmente (sintagmi nominali indefiniti come "un uomo"), la regola è:

Regola per Sintagmi Indefiniti: "Un F è G" Introduce un nuovo referente x Aggiunge le condizioni F(x) e G(x)

Per pronomi anaforici, la regola è:

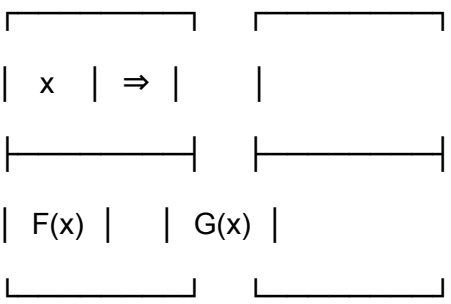
Regola per Pronomi: "Lui/Lei/Esso" Identifica un referente accessibile x nella DRS con le proprietà grammaticali appropriate Non introduce nuovo referente, riutilizza x esistente

Per descrizioni definite, la regola è simile ai pronomi ma con vincoli addizionali di unicità:

Regola per Descrizioni Definite: "Il F" Identifica l'unico referente accessibile x tale che F(x) è già una condizione nella DRS Se non esiste tale referente, introduce nuovo referente e accomoda la presupposizione

Per quantificatori universali come "ogni", la DRT introduce strutture condizionali più complesse:

Regola per "Ogni F è G": Introduce una sotto-DRS condizionale:



Questa rappresentazione cattura l'idea che "ogni F è G" introduce un'implicazione universale: per ogni possibile valore di x che soddisfa F(x), deve essere il caso che G(x).

La semantica formale di una DRS specifica quando una DRS è vera in un modello. Una DRS $K = \langle U\_K, Con\_K \rangle$ è vera in un modello M se e solo se esiste una funzione di embedding $h: U\_K \rightarrow M$ che assegna entità del modello ai referenti discorsivi in modo tale che tutte le condizioni in Con_K siano soddisfatte quando interpretate rispetto a h.

Formalmente:

$M \models K$ sse $\exists h: U\_K \rightarrow Dom(M)$ tale che per ogni condizione $c \in Con\_K$, $M \models c[h]$

dove Dom(M) è il dominio di entità del modello M, e c[h] denota la condizione c interpretata con i referenti discorsivi sostituiti dalle loro assegnazioni secondo h.

Per DRS condizionali della forma $K_1 \Rightarrow K_2$:

$M \models (K_1 \Rightarrow K_2)$ sse per ogni embedding $h_1: U\_K_1 \rightarrow Dom(M)$ tale che $M \models K_1[h_1]$, esiste estensione $h_2: U\_K_1 \cup U\_K_2 \rightarrow Dom(M)$ di h_1 tale che $M \models K_2[h_2]$

Questa definizione cattura formalmente l'idea che una condizione universale richiede che per ogni modo di soddisfare l'antecedente K_1, debba esistere un modo compatibile di soddisfare il conseguente K_2.

Una delle virtù principali della DRT è la sua capacità di trattare sistematicamente fenomeni di anafora cross-sentential che sono problematici per la semantica statica tradizionale. Nella semantica vero-condizionale standard di Montague, ogni frase è interpretata indipendentemente con i suoi quantificatori vincolando variabili soltanto all'interno dello scope sintattico della frase. Non esiste meccanismo per un quantificatore esistenziale in una frase di introdurre una variabile che rimanga accessibile per binding anaforico in frasi successive.

Consideriamo il contrasto seguente nella semantica statica versus dinamica. In logica del primo ordine standard, il seguente discorso sarebbe rappresentato come:

"Un uomo entrò. Lui si sedette." $\exists x$ [Uomo(x) $\wedge$ Entrare(x)] $\wedge$ Sedere(?)

Ma non c'è modo di specificare a cosa il pronome "lui" si riferisce, perché la variabile x introdotta dal quantificatore esistenziale è vincolata soltanto all'interno dello scope di $\exists x$, e non è accessibile nella seconda congiunzione.

La DRT risolve questo problema naturalmente perché i referenti discorsivi introdotti da sintagmi indefiniti non sono vincolati da quantificatori con scope limitato, ma piuttosto sono aggiunti all'universo del discorso U_K della DRS, dove rimangono permanentemente accessibili per riferimento anaforico da parte di frasi successive, finché non vengono esplicitamente rimossi o resi inaccessibili da operatori che creano confini di accessibilità.

Vincoli di Accessibilità e la Gerarchia dei Referenti Discorsivi

Non tutti i referenti discorsivi introdotti nel corso di un discorso rimangono uniformemente accessibili per riferimento anaforico successivo. Esistono vincoli strutturali e pragmatici sull'accessibilità che determinano quali referenti possono essere legittimamente ripresi da espressioni anaforiche in punti specifici del discorso.

Un vincolo strutturale fondamentale riguarda le condizioni subordinate e i contesti modali o ipotetici. Consideriamo:

(36) Se un gatto$_1$ entra, (lui$_1$) miagola. (37) *Un gatto$_1$ potrebbe entrare. Lui$_1$ miagola.

In (36), il pronome "lui" può legittimamente riferirsi al gatto introdotto nell'antecedente del condizionale, anche se questo referente è introdotto in un contesto ipotetico subordinato. Tuttavia, il riferimento anaforico è limitato allo scope del condizionale.

In (37), il tentativo di riferimento anaforico al gatto introdotto nel contesto modale "potrebbe" fallisce quando si esce da quel contesto modale. Il secondo enunciato non può essere interpretato come riferentesi allo stesso gatto possibile introdotto dal primo enunciato.

Questo contrasto può essere spiegato nella DRT attraverso la nozione di accessibilità subordinata. I referenti introdotti all'interno di sub-DRS subordinate (create da condizionali, modali, negazioni, etc.) non sono accessibili al di fuori di quelle sub-DRS. Formalmente, definiamo una relazione di accessibilità tra referenti e posizioni nel discorso:

Un referente x è accessibile alla posizione p nel discorso sse x è introdotto in una DRS K tale che K domina la DRS corrispondente alla posizione p nella struttura gerarchica delle DRS.

La relazione di dominanza tra DRS è definita ricorsivamente: una DRS K_1 domina una DRS K_2 se K_2 è una sub-DRS diretta di K_1, o se K_1 domina una DRS K_3 che a sua volta domina K_2 (dominanza transitiva). Inoltre, una DRS domina sempre se stessa (dominanza riflessiva).

Per esempio, nella rappresentazione DRT di un condizionale "Se φ allora ψ", i referenti introdotti nell'antecedente φ sono accessibili nel conseguente ψ (perché entrambi sono dominati dalla DRS condizionale complessiva), ma i referenti introdotti nel conseguente non sono accessibili al di fuori del condizionale (perché la DRS del conseguente non domina la DRS del discorso circostante).

Oltre ai vincoli strutturali, esistono anche vincoli pragmatici sull'accessibilità basati sulla salienza discorsiva e sulla centratura (centering). La teoria della centratura (Centering Theory), sviluppata da Barbara Grosz, Aravind Joshi e Scott Weinstein negli anni Novanta, propone che in ogni punto del discorso, alcuni referenti sono più centrali o salienti di altri, e che le espressioni anaforiche tendono preferenzialmente a riferirsi a questi referenti centrali.

La teoria postula che ogni enunciato in un discorso è associato a due liste ordinate di referenti:

1. Forward-looking centers (Cf): la lista di tutti i referenti menzionati nell'enunciato corrente, ordinati per salienza (tipicamente determinata da fattori come ruolo grammaticale, con i soggetti essendo più salienti degli oggetti, ordine di menzione, animatezza, etc.)

2. Backward-looking center (Cb): il referente più saliente dell'enunciato precedente che è anche menzionato nell'enunciato corrente, ovvero il referente che viene "mantenuto" attraverso la transizione discorsiva.

La teoria della centratura formula vincoli sulla coerenza locale del discorso in termini di transizioni tra stati di centratura. Le transizioni più coerenti sono quelle che mantengono lo stesso backward-looking center attraverso enunciati successivi (CONTINUAZIONE), seguite da transizioni che mantengono il centro ma lo retrogradano nella lista dei forward-looking centers (RITENZIONE), seguite da transizioni che cambiano il centro (SHIFT).

Formalmente, data una sequenza di enunciati U_1, U_2, ..., U_n, definiamo:

Cf(Ui) = lista ordinata dei referenti menzionati in Ui Cb(Ui) = il referente più saliente in Cf(Ui-1) che appare anche in Cf(Ui)

Le transizioni sono classificate come:

CONTINUAZIONE: Cb(Ui) = Cb(Ui-1) e Cb(Ui) = elemento più saliente in Cf(Ui) RITENZIONE: Cb(Ui) = Cb(Ui-1) e Cb(Ui) ≠ elemento più saliente in Cf(Ui) SMOOTH-SHIFT:

Cb(Ui) ≠ Cb(Ui-1) e Cb(Ui) = elemento più saliente in Cf(Ui) ROUGH-SHIFT: Cb(Ui) ≠ Cb(Ui-1) e Cb(Ui) ≠ elemento più saliente in Cf(Ui)

La preferenza di coerenza è: CONTINUAZIONE > RITENZIONE > SMOOTH-SHIFT > ROUGH-SHIFT

Questa gerarchia predice che i discorsi che mantengono lo stesso centro attraverso transizioni successive sono percepiti come più coerenti e più facili da processare di quelli che cambiano frequentemente centro, una predizione che è stata confermata sperimentalmente attraverso studi di comprensione che misurano tempi di lettura e carichi cognitivi.

La File Change Semantics: Formalizzazione Completa del Potenziale di Cambiamento del Contesto

Proseguendo con la formalizzazione rigorosa della File Change Semantics, è essenziale specificare le regole per tutti i costrutti logici fondamentali per completare il sistema. Per la negazione, Heim propone una regola che differisce radicalmente dal trattamento standard nella logica classica:

Regola per Negazione (prima versione): $\underline{\neg\varphi} = F \setminus \underline{\varphi} = \{g \in F \mid g \notin \underline{\varphi}\}$

Questa regola definisce la negazione come una sorta di "test": il file viene aggiornato con ¬φ rimuovendo tutte le assegnazioni che sarebbero state mantenute dall'aggiornamento con φ. Intuitivamente, ¬φ elimina dal contesto tutte le possibilità che renderebbero φ vero.

Tuttavia, questa regola per la negazione ha una conseguenza teorica importante: la negazione diventa un operatore non-persistente che non introduce permanentemente informazione negativa nel file. Consideriamo la sequenza di aggiornamenti:

$F_0 \rightarrow \underline{\varphi} = F_1 \rightarrow \underline{\neg\varphi} = F_1 \setminus F_1 = \varnothing$

Se partiamo da un file F_0 e lo aggiorniamo prima con φ e poi con ¬φ, otteniamo il file vuoto, il che rappresenta uno stato informazionale contraddittorio dove nessuna possibilità è compatibile con l'informazione stabilita. Questo cattura correttamente l'intuizione che asserire φ e poi ¬φ genera contraddizione.

Ma consideriamo la sequenza inversa:

$F_0 \rightarrow \underline{\neg\varphi} = F_0 \setminus \underline{\varphi} = F_1 \rightarrow \underline{\varphi} = F_1 \cap \underline{\varphi}$

Qui il risultato non è necessariamente il file vuoto. Se prima eliminiamo le possibilità che renderebbero φ vero, e poi cerchiamo di aggiungere φ, otteniamo soltanto quelle possibilità che erano nel file ristretto F_1 e che soddisfano φ, il che potrebbe essere vuoto ma non lo è necessariamente se il file originale F_0 non determinava completamente la verità o falsità di φ.

Questa asimmetria riflette il fatto che la negazione nella FCS è "debole" nel senso che non aggiunge positivamente informazione ma soltanto elimina possibilità. Per ottenere una semantica più standard dove ¬¬φ è equivalente a φ, potremmo adottare una regola alternativa:

Regola per Negazione (versione forte): $\underline{\neg\varphi} = \{g \in F \mid \forall g' \text{ compatibile con } g, g' \notin \underline{\varphi}\}$

dove F_maximal è il file massimale che contiene tutte le possibili funzioni di assegnazione. Questa versione stabilisce che ¬φ mantiene soltanto quelle assegnazioni che sono incompatibili con φ nel senso più forte possibile.

Per i quantificatori universali, la FCS adotta una strategia che può sembrare inizialmente contro-intuitiva ma che si rivela empiricamente adeguata per catturare i pattern di anafora:

Regola per Quantificazione Universale: $\underline{\forall x\, \varphi} = F$ se per ogni $g \in F$ e ogni $d \in D$, $g[x/d] \in [\varphi](F \cup \{g[x/d]\}) = \varnothing$ altrimenti

Questa regola verifica che per ogni modo di estendere le assegnazioni nel file corrente con un valore per x, quella estensione risulterebbe in un file che soddisfa φ. Se questa condizione è soddisfatta, il file passa attraverso invariato (test riuscito). Se fallisce, il file diventa vuoto (test fallito, contraddizione).

Crucialmente, il quantificatore universale non introduce permanentemente il referente discorsivo x nel file, a differenza del quantificatore esistenziale. Questo spiega l'asimmetria nell'accessibilità anaforica che abbiamo osservato precedentemente: i referenti introdotti esistenzialmente rimangono accessibili per anafora successiva, mentre quelli introdotti universalmente no.

Un aspetto particolarmente elegante della FCS è come tratta le descrizioni definite, che avevamo analizzato estensivamente nel contesto della teoria di Russell. Nella FCS, una descrizione definita "il F" presuppone che esiste esattamente un individuo che soddisfa F nel file corrente, e asserisce che quell'individuo soddisfa qualsiasi predicato a cui la descrizione è applicata.

Regola per Descrizioni Definite: $[\underline{\text{il } x\text{: } F(x)}](F)$ è definita solo se esiste esattamente un $d \in D$ tale che per ogni $g \in F$, $g[x/d] \in [F(x)](F \cup \{g[x/d]\})$ Se definita, allora $[\underline{\text{il } x\text{: } F(x)}](F) = [G(x)](F \cup \{g[x/d] \mid g \in F\})$ dove d è l'unico individuo che soddisfa la condizione sopra.

Questa regola incorpora naturalmente il trattamento presupposizionale delle descrizioni definite: se la presupposizione di unicità ed esistenza non è soddisfatta nel file corrente, la funzione di cambiamento contesto è indefinita, rappresentando il fallimento presupposizionale. Se la presupposizione è soddisfatta, il file viene aggiornato con l'informazione che l'individuo unico che soddisfa F anche soddisfa G.

La FCS fornisce anche un meccanismo elegante per modellare l'accomodamento presupposizionale che avevamo discusso precedentemente. Quando una descrizione definita viene usata ma la sua presupposizione non è soddisfatta nel file corrente, l'ascoltatore può accomodare la presupposizione restringendo tacitamente il file per

includere soltanto quelle assegnazioni che soddisfano la presupposizione esistenziale e di unicità. Formalmente:

Accomodamento Presupposizionale: Se [il x: F(x)](F) è indefinita perché la presupposizione non è soddisfatta, allora accomoda restringendo F a F' = {g ∈ F | ∃!d [g[x/d] ∈ F(x)]} e applica [il x: F(x)] a F'.

Questo meccanismo permette all'ascoltatore di "aggiustare" il contesto per rendere felicitous un enunciato che altrimenti sarebbe presupposizionalmente difettoso, a patto che l'aggiustamento sia compatibile con l'informazione già stabilita e con la conoscenza del mondo dell'ascoltatore.

L'Architettura Neurale della Comprensione Linguistica: Dalla Semantica Formale alla Neurobiologia

Avendo esplorato in profondità i fondamenti formali della semantica compositiva, dinamica e pragmatica, è ora essenziale connettere queste teorie astratte con le realtà neurobiologiche concrete del cervello umano che implementa effettivamente questi processi computazionali. Come vengono realizzati, a livello di circuiti neurali e dinamiche neurofisiologiche, i processi di analisi sintattica, interpretazione semantica, arricchimento pragmatico, e integrazione discorsiva che le teorie formali descrivono in termini di manipolazione di strutture simboliche astratte?

Le neuroscienze cognitive del linguaggio hanno compiuto progressi straordinari nelle ultime tre decadi grazie alla convergenza di molteplici metodologie sperimentali complementari: la neuroimaging funzionale attraverso risonanza magnetica funzionale (fMRI) che rivela quali regioni cerebrali sono attive durante specifici compiti linguistici con risoluzione spaziale dell'ordine del millimetro ma risoluzione temporale limitata a secondi; la magnetoencefalografia (MEG) e l'elettroencefalografia (EEG) che misurano l'attività elettromagnetica cerebrale con risoluzione temporale del millisecondo ma risoluzione spaziale più limitata; la stimolazione magnetica transcranica (TMS) che permette di interferire temporaneamente con l'attività di regioni corticali specifiche per testare la loro necessità causale per determinate funzioni; e gli studi su pazienti con lesioni cerebrali focali che rivelano le conseguenze comportamentali del danneggiamento di specifiche strutture neurali.

La convergenza di evidenze da queste metodologie diverse ha portato a una revisione sostanziale del modello classico di Wernicke-Lichtheim-Geschwind che dominava la neurologia del linguaggio per più di un secolo. Il modello classico, nella sua formulazione più schematica, postulava una divisione del lavoro relativamente netta tra l'area di Broca nella circonvoluzione frontale inferiore sinistra (aree di Brodmann quarantaquattro e quarantacinque), responsabile primariamente della produzione linguistica e dell'elaborazione sintattica, e l'area di Wernicke nel giro temporale superiore posteriore sinistro (parte dell'area ventidue di Brodmann), responsabile primariamente della comprensione linguistica e dell'elaborazione semantica, con il fascicolo arcuato che fornisce la connessione anatomica tra queste due regioni permettendo la ripetizione di materiale linguistico ascoltato.

Questo modello, sebbene storicamente importante e ancora pedagogicamente utile come prima approssimazione, si è rivelato drammaticamente semplificato alla luce delle evidenze contemporanee. Gli studi di neuroimaging funzionale rivelano che l'elaborazione linguistica coinvolge reti estese e distribuite che si estendono ben oltre le aree di Broca e Wernicke classiche, includendo regioni nella corteccia temporale media e inferiore, nel giro angolare e sopramarginale del lobo parietale inferiore, nella corteccia prefrontale dorsolaterale e ventromediale, e persino in regioni sottocorticali come il talamo, i gangli della base, e il cervelletto.

Inoltre, la dicotomia sintassi-semantica associata con Broca versus Wernicke si è rivelata eccessivamente schematica. Entrambe le regioni sono coinvolte sia nell'elaborazione sintattica sia in quella semantica, sebbene con contributi funzionali differenziati. Gli studi fMRI che manipolano sistematicamente la complessità sintattica (per esempio confrontando frasi con embedding multiplo versus frasi sintatticamente semplici) rivelano attivazione non soltanto nell'area di Broca ma anche in regioni temporali posteriori e parietali. Similmente, compiti che richiedono elaborazione semantica intensiva attivano non soltanto regioni temporali ma anche regioni frontali inferiori.

Un modello contemporaneo più sofisticato è il modello dual-stream proposto da Gregory Hickok e David Poeppel, che postula l'esistenza di due vie (streams) di elaborazione parallele analoghe alle vie "what" e "where" nel sistema visivo. La via ventrale (ventral stream) proietta dalla corteccia uditiva primaria nel giro di Heschl attraverso il giro temporale superiore posteriore verso il giro temporale medio e il polo temporale anteriore, ed è responsabile primariamente per la mappatura suono-a-significato, ovvero l'elaborazione semantica lessicale e compositiva che trasforma rappresentazioni fonologiche in rappresentazioni semantiche. La via dorsale (dorsal stream) proietta dalla corteccia uditiva primaria attraverso il giro sopramarginale nel lobo parietale inferiore verso la corteccia premotoria ventrale includendo l'area di Broca, ed è responsabile primariamente per la mappatura suono-a-articolazione, ovvero i processi che supportano la ripetizione, la working memory fonologica, e il controllo motorio della produzione linguistica.

Formalmente, possiamo concepire la via ventrale come implementante una funzione di mappatura:

f_ventrale: Rappresentazioni_Fonologiche → Rappresentazioni_Semantiche

mentre la via dorsale implementa una funzione di mappatura:

f_dorsale: Rappresentazioni_Fonologiche → Rappresentazioni_Motorie_Articolatorie

Queste due vie operano in parallelo durante la comprensione linguistica normale, con la via ventrale che estrae il significato dall'input acustico e la via dorsale che mantiene rappresentazioni fonologiche in working memory e supporta potenzialmente processi di simulazione articolatoria implicita che facilitano la percezione del parlato in condizioni di rumore o degradazione del segnale.

Le evidenze a supporto di questo modello dual-stream provengono da molteplici fonti convergenti. Gli studi di neuroimaging funzionale rivelano che compiti semantici (come giudizi di sinonimia, categorizzazione semantica, o generazione di associati semantici)

attivano preferenzialmente regioni della via ventrale nel giro temporale medio e inferiore, mentre compiti fonologici (come giudizi di rima, segmentazione fonemica, o ripetizione di non-parole) attivano preferenzialmente regioni della via dorsale nel giro sopramarginale e nella corteccia premotoria. Gli studi su pazienti con lesioni forniscono evidenza ancora più diretta: pazienti con lesioni dorsali (per esempio nel giro sopramarginale) manifestano deficit selettivi nella ripetizione e nella working memory fonologica ma comprensione semantica relativamente preservata, mentre pazienti con lesioni ventrali (per esempio nel giro temporale medio posteriore) manifestano deficit selettivi nella comprensione semantica ma capacità di ripetizione relativamente preservata.

A un livello di dettaglio ancora maggiore, possiamo considerare la dinamica temporale fine-grained dell'elaborazione linguistica rivelata dagli studi di potenziali evento-correlati (ERP). Come avevamo menzionato precedentemente, gli studi ERP hanno identificato diverse componenti che riflettono stadi specifici di elaborazione linguistica con risoluzione temporale del millisecondo.

La componente N400, una deflessione negativa del potenziale elettrico che raggiunge il picco approssimativamente quattrocento millisecondi dopo la presentazione di una parola, è stata associata consistentemente attraverso letteralmente migliaia di studi con l'elaborazione semantica lessicale e l'integrazione semantica contestuale. L'ampiezza della N400 è modulata da molteplici fattori che influenzano la facilità di accesso e integrazione semantica:

1. **Prevedibilità Contestuale**: Parole altamente prevedibili dal contesto precedente evocano una N400 di ampiezza ridotta rispetto a parole imprevedibili. Per esempio, nel contesto "Il bambino giocava con il suo _____", la parola "giocattolo" evoca una N400 piccola, mentre la parola "lavandino" evoca una N400 grande.

2. **Frequenza Lessicale**: Parole ad alta frequenza nella lingua evocano N400 ridotte rispetto a parole a bassa frequenza, riflettendo la facilitazione dell'accesso lessicale per items più familiari.

3. **Priming Semantico**: Parole semanticamente correlate al contesto precedente evocano N400 ridotte. Per esempio, "infermiera" evoca N400 ridotta dopo "dottore" rispetto a dopo "tavolo".

4. **Congruenza Semantica**: Violazioni di selezione semantica o incongruenze semantiche palesi evocano N400 di ampiezza molto grande. L'esempio classico è: "Ho messo lo zucchero nel caffè" (N400 piccola su "zucchero") versus "Ho messo i calzini nel caffè" (N400 grande su "calzini").

L'interpretazione funzionale della N400 che ha ricevuto il supporto empirico più consistente è che riflette la difficoltà o lo sforzo computazionale coinvolto nell'integrare la rappresentazione semantica della parola corrente con la rappresentazione semantica del contesto precedente. Formalmente, possiamo modellare l'ampiezza della N400 come una funzione della distanza semantica tra la parola corrente e il contesto:

Ampiezza_N400(parola, contesto) = k · Distanza_Semantica([[parola]], [[contesto]]) + ε

dove [[parola]] e [[contesto]] sono rappresentazioni semantiche (potenzialmente vettori in uno spazio semantico multi-dimensionale), Distanza_Semantica è una metrica (per esempio distanza euclidea o distanza del coseno) che quantifica quanto la rappresentazione della parola è distante o incongruente con quella del contesto, k è una costante di proporzionalità, e ε rappresenta noise neurale e variabilità individuale.

Questa formalizzazione predice quantitativamente che quanto più una parola è semanticamente distante o incongruente con il contesto, tanto maggiore sarà l'ampiezza della N400 evocata, una predizione che è stata confermata ripetutamente attraverso studi che manipolano sistematicamente la similarità semantica tra parole target e contesti.

Le fonti neurali della N400 sono state localizzate attraverso tecniche di localizzazione della sorgente EEG/MEG e studi intracranici su pazienti epilettici con elettrodi impiantati. Le evidenze convergono nell'identificare un network distribuito di regioni che contribuiscono alla generazione della N400, includendo:

- Il giro temporale medio bilateralmente, con predominanza nell'emisfero sinistro
- Il giro temporale inferiore
- Il giro frontale inferiore (area di Broca)
- Regioni prefrontali mediali

Questa distribuzione spaziale è consistente con l'idea che la N400 riflette un processo di integrazione semantica che coinvolge molteplici componenti del sistema semantico: l'accesso a rappresentazioni semantiche lessicali nel giro temporale medio, il recupero di conoscenza concettuale dal giro temporale inferiore, e processi di selezione e controllo esecutivo dalle regioni prefrontali.

La componente P600, una deflessione positiva che emerge tipicamente tra cinquecento e ottocento millisecondi dopo la presentazione dello stimolo critico, è stata associata classicamente con l'elaborazione sintattica, particolarmente con processi di rianalisi sintattica, riparazione di violazioni sintattiche, e integrazione di strutture sintattiche complesse. La P600 viene evocata da:

1. **Violazioni Sintattiche Manifeste**: Errori di accordo numero, genere, caso; ordine delle parole agrammaticale; violazioni di subcategorizzazione verbale. Per esempio: "*Il libro leggono interessante" evoca P600 grande su "leggono" che viola l'accordo soggetto-verbo.

2. **Frasi Garden-Path**: Costruzioni sintatticamente ambigue che inducono il parser a costruire inizialmente un'analisi errata che deve essere successivamente rianalizzata. Per esempio: "Il cavallo cavalcato dal fantino cadde" evoca P600 su "cavalcato" quando il parser realizza che "cavallo" è paziente, non agente.

3. **Complessità Sintattica**: Dipendenze a lunga distanza, embedding multiplo, e altre strutture sintatticamente complesse possono evocare P600 anche in assenza di violazioni o ambiguità, riflettendo il maggior carico computazionale dell'integrazione sintattica.

L'interpretazione funzionale della P600 è stata oggetto di dibattito teorico sostanziale. L'interpretazione classica la associa con processi di rianalisi e riparazione sintattica: quando il parser incontra evidenza che la sua analisi sintattica corrente è incompatibile con l'input, la P600 riflette il processo di costruzione di un'analisi alternativa. Formalmente:

$$P600 \propto Costo\_Computazionale(Rianalisi\_Sintattica)$$

dove il costo computazionale potrebbe essere quantificato in termini del numero di operazioni di ristrutturazione necessarie, della profondità nell'albero sintattico dove la rianalisi deve avvenire, o di altri fattori strutturali.

Un'interpretazione alternativa più recente, proposta da Gina Kuperberg e collaboratori, suggerisce che la P600 riflette un conflitto tra due vie di elaborazione parallele: una via combinatoriale che costruisce rappresentazioni sintattiche compositive in modo bottom-up, e una via basata su memoria che genera aspettative semantiche basate su associazioni lessicali e schemi di eventi probabilistici. Quando queste due vie producono interpretazioni incompatibili, emerge la P600 riflettendo il processo di risoluzione del conflitto e riconciliazione delle rappresentazioni discordanti.

Questa interpretazione basata sul conflitto può spiegare alcuni pattern empirici che sono problematici per l'interpretazione puramente sintattica della P600, in particolare il fenomeno delle "semantic P600s" osservate in frasi come:

(20) Il uova furono nel frigo dal ragazzo. [violazione semantica evoca P600, non N400]

Qui la violazione di selezione semantica (le uova non possono agentivamente mettere qualcosa nel frigo) evoca una P600 piuttosto che la N400 che ci aspetteremmo per una violazione semantica. Secondo la teoria del conflitto, questo riflette il fatto che le forti associazioni lessicali tra "uova" e "frigo" generate dalla via basata su memoria entrano in conflitto con l'interpretazione combinatoria che assegna a "uova" il ruolo di agente.

Modelli Computazionali Neurali della Sintassi e Semantica Compositiva

Proseguendo con l'analisi dell'implementazione neurale dell'elaborazione linguistica, è fondamentale considerare come le operazioni compositive formali che abbiamo analizzato estensivamente nella semantica di Montague, nella DRT, e nella FCS possano essere realizzate attraverso circuiti neurali biologici. Questa domanda costituisce uno dei problemi più profondi nelle neuroscienze cognitive contemporanee: il problema del binding, ovvero come il cervello rappresenta e manipola strutture composizionali gerarchiche usando reti neurali che sono fondamentalmente parallele e distribuite.

Il problema può essere illustrato considerando una frase semplice come "Il gatto insegue il topo". La rappresentazione semantica compositiva di questa frase richiede di rappresentare non soltanto i concetti individuali GATTO, INSEGUIRE, TOPO, ma anche le relazioni strutturali tra questi concetti: quale entità è l'agente dell'azione di inseguire e quale è il paziente. Nella rappresentazione logica formale, scriviamo:

Insegue(il_gatto, il_topo)

dove la struttura sintattica della formula specifica inequivocabilmente che il_gatto occupa la posizione di argomento esterno (agente) e il_topo occupa la posizione di argomento interno (paziente). Ma come può un network neurale, che consiste di neuroni che codificano informazione attraverso i loro tassi di firing e le loro correlazioni temporali, rappresentare questa struttura proposizionale composizionale?

Se rappresentiamo semplicemente i concetti attraverso l'attivazione di popolazioni neuronali distinte, senza alcun meccanismo per codificare le relazioni strutturali tra di essi, allora il network non potrebbe distinguere "Il gatto insegue il topo" da "Il topo insegue il gatto", perché le stesse popolazioni neuronali sarebbero attive in entrambi i casi. Questo è il problema del binding variabile (variable binding problem): come possiamo implementare neuralmente la distinzione tra Insegue(x, y) e Insegue(y, x) quando i ruoli di x e y sono invertiti?

Diverse soluzioni computazionali sono state proposte nella letteratura delle neuroscienze computazionali e dell'intelligenza artificiale neuro-simbolica. Una classe di soluzioni si basa sulla sincronizzazione temporale (temporal synchrony binding): l'idea è che neuroni che codificano elementi che devono essere legati insieme (bound together) nella stessa struttura composizionale scaricano sincronicamente, ovvero con le loro spike che sono temporalmente allineate entro finestre di pochi millisecondi. Secondo questa teoria, quando processiamo "Il gatto insegue il topo", le popolazioni neuronali che rappresentano GATTO e AGENTE scaricherebbero sincronicamente, così come le popolazioni che rappresentano TOPO e PAZIENTE, ma ci sarebbe decorrelazione temporale tra i neuroni che rappresentano GATTO e quelli che rappresentano PAZIENTE.

Formalmente, possiamo modellare questo attraverso una funzione di correlazione temporale. Sia $n_i(t)$ il tasso di firing del neurone i al tempo t. La correlazione temporale tra neuroni i e j in una finestra temporale $[t, t+\Delta t]$ è:

$$C_{ij}(t) = (1/\Delta t) \int_{[t \text{ to } t+\Delta t]} n_i(\tau) \cdot n_j(\tau - \delta)\, d\tau$$

dove δ è un lag temporale piccolo. Se C_{ij} è alto, i neuroni i e j sono altamente sincronizzati. Il binding composizionale sarebbe implementato attraverso pattern di sincronizzazione:

Rappresentazione("Il gatto insegue il topo"):

- Cluster sincrono 1: {neuroni_GATTO, neuroni_AGENTE} con alta correlazione interna
- Cluster sincrono 2: {neuroni_TOPO, neuroni_PAZIENTE} con alta correlazione interna
- Bassa correlazione tra cluster 1 e cluster 2

Rappresentazione("Il topo insegue il gatto"):

- Cluster sincrono 1: {neuroni_TOPO, neuroni_AGENTE}
- Cluster sincrono 2: {neuroni_GATTO, neuroni_PAZIENTE}

Questa architettura permette di rappresentare distintamente le due frasi usando le stesse popolazioni neuronali sottostanti ma con pattern di sincronizzazione differenti.

Le evidenze empiriche per il temporal synchrony binding nel linguaggio sono promettenti ma ancora non conclusive. Gli studi MEG che misurano l'attività cerebrale oscillatoria durante la comprensione di frasi rivelano che l'elaborazione sintattica è associata con aumenti della coerenza di fase nella banda theta (quattro-otto Hertz) e gamma (trenta-ottanta Hertz) tra regioni frontali e temporali. Peter Hagoort e collaboratori hanno mostrato che la coerenza theta aumenta durante l'elaborazione di frasi sintatticamente complesse rispetto a liste di parole, e che questa sincronizzazione a lunga distanza è necessaria per l'integrazione sintattica: la stimolazione magnetica transcranica ritmica (rhythmic TMS) che perturba le oscillazioni theta nella corteccia frontale inferiore sinistra interferisce selettivamente con la comprensione di frasi sintatticamente complesse.

Una soluzione alternativa al problema del binding è fornita dalle architetture tensor product, sviluppate da Paul Smolensky e collaboratori. In questa architettura, le rappresentazioni composizionali sono codificate come prodotti tensoriali di rappresentazioni di ruoli (roles) e riempitori (fillers).

Formalmente, sia f un vettore che rappresenta un filler (per esempio, il concetto GATTO), e sia r un vettore che rappresenta un ruolo (per esempio, AGENTE). Il binding di f a r è rappresentato dal prodotto tensoriale f ⊗ r, che è una matrice (o più generalmente, un tensore di ordine superiore) i cui elementi sono dati da:

$(f \otimes r)_{ij} = f_i \cdot r_j$

Una rappresentazione composizionale complessa che coinvolge molteplici bindings può essere costruita sommando i prodotti tensoriali individuali:

Rappresentazione("Il gatto insegue il topo") = GATTO ⊗ AGENTE + TOPO ⊗ PAZIENTE + INSEGUIRE ⊗ PREDICATO

Questa rappresentazione tensoriale preserva tutta l'informazione composizionale: possiamo recuperare quale concetto occupa quale ruolo mediante operazioni di unbinding che proiettano la rappresentazione tensoriale su specifici sottospazi vettoriali corrispondenti ai ruoli. Per estrarre il filler del ruolo AGENTE:

Filler_AGENTE = (GATTO ⊗ AGENTE + TOPO ⊗ PAZIENTE + INSEGUIRE ⊗ PREDICATO) · AGENTE^T = GATTO · (AGENTE · AGENTE^T) + TOPO · (PAZIENTE · AGENTE^T) + INSEGUIRE · (PREDICATO · AGENTE^T) ≈ GATTO (se i vettori di ruolo sono approssimativamente ortogonali)

dove AGENTE^T è il vettore di ruolo trasposto usato come probe per estrarre il filler.

L'architettura tensor product può essere implementata neuralmente attraverso reti che computano prodotti tensoriali distribuiti. Ogni neurone nella rete potrebbe codificare l'attivazione di un particolare elemento del tensore, ovvero una particolare combinazione di features del filler e features del ruolo. La popolazione neuronale complessiva rappresenta così distributamente l'intera struttura compositiva.

Recentemente, i modelli di linguaggio neurali basati su architetture Transformer, particolarmente modelli come BERT e GPT che abbiamo menzionato precedentemente, hanno dimostrato capacità sorprendenti di catturare strutture sintattiche e semantiche composizionali nonostante siano addestrati esclusivamente attraverso obiettivi di predizione linguistica auto-supervisionati, senza alcuna supervisione sintattica esplicita. Questi modelli utilizzano meccanismi di attenzione multi-testa che permettono a ogni posizione nella sequenza di input di attendere selettivamente a tutte le altre posizioni, creando rappresentazioni contestualizzate che codificano informazione circa le dipendenze strutturali tra parole.

Formalmente, il meccanismo di self-attention in un Transformer computa, per ogni token nella posizione i, una rappresentazione contestualizzata che è una combinazione pesata delle rappresentazioni di tutti i token:

$h_i = \Sigma_j \alpha_{ij} \cdot v_j$

dove v_j sono i "value vectors" associati a ciascun token j, e α_{ij} sono pesi di attenzione che determinano quanto il token in posizione i dovrebbe attendere al token in posizione j. I pesi di attenzione sono computati attraverso un meccanismo di soft-max su similarity scores:

$\alpha_{ij} = \exp(\text{score}(q_i, k_j)) / \Sigma_k \exp(\text{score}(q_i, k_k))$

dove q_i è il "query vector" della posizione i, k_j è il "key vector" della posizione j, e score è tipicamente il prodotto scalare:

$\text{score}(q_i, k_j) = (q_i \cdot k_j) / \sqrt{d_k}$

con d_k la dimensionalità dei vettori chiave, e la divisione per $\sqrt{d_k}$ che fornisce normalizzazione per stabilità numerica.

Analisi recenti delle rappresentazioni interne apprese da questi modelli rivelano che diverse teste di attenzione (attention heads) nelle architetture multi-testa si specializzano spontaneamente per catturare diversi tipi di relazioni sintattiche. Per esempio, Kevin Clark e collaboratori hanno mostrato che in BERT alcune teste di attenzione attendono sistematicamente da oggetti diretti ai loro verbi governanti, altre da soggetti ai verbi corrispondenti, altre da determinanti ai nomi che modificano, e così via, ricapitolando sostanzialmente l'architettura delle dipendenze sintattiche anche se il modello non è stato mai addestrato esplicitamente su alberi sintattici annotati.

Questa emergenza spontanea di sensibilità sintattica da obiettivi puramente predittivi solleva questioni profonde circa la natura della conoscenza sintattica e il suo ruolo funzionale nella comprensione linguistica. Forse la sintassi non è un modulo cognitivo autonomo con rappresentazioni e principi propri indipendenti dalla semantica e dalla pragmatica, come sostenuto dalle teorie generativiste classiche, ma piuttosto emerge come una conseguenza statistica dell'ottimizzazione per la predizione e l'integrazione di informazione semantica attraverso contesti linguistici. Questa prospettiva è più compatibile con le teorie costruzioniste e usage-based dell'acquisizione linguistica che enfatizzano il ruolo dell'apprendimento statistico e della generalizzazione basata su pattern.

L'Acquisizione del Linguaggio: Innatismo versus Empirismo Revisionato

La questione dell'acquisizione del linguaggio costituisce uno dei dibattiti teorici più fondamentali e duraturi nelle scienze cognitive, toccando questioni profonde circa la natura della mente umana, il ruolo dell'esperienza versus dotazione genetica nello sviluppo cognitivo, e la specificità versus generalità dei meccanismi di apprendimento.

La posizione innatista, articolata classicamente e difesa vigorosamente da Noam Chomsky e i suoi seguaci nella tradizione generativista, si basa sull'argomento della povertà dello stimolo: l'osservazione che i bambini sono esposti a un input linguistico quantitativamente limitato, qualitativamente degradato (contenente errori, false partenze, enunciati incompleti), e soprattutto privo di evidenza negativa sistematica (feedback esplicito su cosa è agrammaticale), eppure convergono rapidamente e uniformemente su grammatiche che incorporano proprietà strutturali complesse e astratte che vanno ben oltre ciò che è esplicitamente manifestato nell'input osservabile.

Possiamo formalizzare l'argomento della povertà dello stimolo nel modo seguente. Sia H lo spazio delle ipotesi grammaticali possibili che un bambino potrebbe in principio intrattenere, ovvero l'insieme di tutte le grammatiche logicamente coerenti. Sia D il Primary Linguistic Data disponibile al bambino, ovvero l'insieme di enunciati linguistici a cui il bambino è stato esposto durante i primi anni di vita. Il problema dell'acquisizione linguistica è trovare la grammatica corretta $G\_target \in H$ basandosi soltanto su D.

Se non ci sono vincoli innati su H, allora $|H|$ è vastissimo, potenzialmente infinito. Per qualsiasi insieme finito di dati D, esistono infinitamente molte grammatiche $G \in H$ che sono compatibili con D nel senso che generano tutti gli enunciati in D, ma che differiscono arbitrariamente nelle loro predizioni su enunciati non contenuti in D. Questo è il problema classico dell'induzione: non possiamo inferire una regola generale da un numero finito di esempi senza assunzioni prior che vincolano lo spazio delle ipotesi.

Formalmente, date osservazioni $D = \{e_1, e_2, ..., e\_n\}$ dove ogni e_i è un enunciato linguistico, lo spazio delle grammatiche compatibili è:

$$Compatible(D) = \{G \in H \mid \forall e\_i \in D, G \text{ genera } e\_i\}$$

Senza vincoli innati, $|Compatible(D)|$ è tipicamente infinito per qualsiasi D finito, rendendo l'apprendimento sottodeterminato.

La soluzione chomskiana è postulare che H è drasticamente ristretto da vincoli innati incorporati nella dotazione genetica umana come parte della Grammatica Universale (UG). UG specifica un insieme di principi universali che devono essere soddisfatti da qualsiasi grammatica umana possibile, e un insieme di parametri binari o finiti che possono variare tra lingue. Lo spazio delle grammatiche è quindi:

$$H\_UG = \{G \mid G \text{ soddisfa tutti i principi di UG}\}$$

dove |H_UG| << |H|. Idealmente, l'esposizione a un piccolo insieme di dati critici D permetterebbe al bambino di fissare i valori parametrici e identificare univocamente G_target.

Le evidenze empiriche che sono state invocate a supporto dell'innatismo includono:

1. Uniformità cross-linguistica nell'acquisizione: I bambini in culture diverse, esposti a lingue tipologicamente diverse, attraversano stadi di sviluppo linguistico notevolmente simili con tempistiche comparabili. La prima parola emerge tipicamente intorno ai dodici mesi, la fase delle combinazioni di due parole intorno ai diciotto-ventiquattro mesi, e le costruzioni sintattiche complesse emergono tra i tre e i quattro anni. Questa uniformità suggerisce che il processo è guidato da un programma maturazionale biologico piuttosto che da fattori ambientali culturalmente variabili.

2. Rapidità dell'acquisizione: I bambini acquisiscono competenza linguistica sofisticata straordinariamente rapidamente, tipicamente padroneggiando la grammatica della loro lingua nativa entro i cinque anni nonostante la complessità strutturale immensa del sistema linguistico. Questa rapidità sembra difficile da riconciliare con meccanismi di apprendimento generale-purpose che richiederebbero presumibilmente molto più tempo e molto più dati.

3. Convergenza su proprietà non osservabili: I bambini convergono su generalizzazioni grammaticali che governano costruzioni rare o assenti nell'input. Per esempio, i bambini anglofoni acquisiscono la conoscenza che le domande si formano attraverso movimento structure-dependent piuttosto che linear, anche se la maggior parte delle frasi nell'input sono troppo semplici per discriminare tra queste due ipotesi (come discusso precedentemente con l'esempio delle domande con clausole relative).

4. Assenza di apprendimento da evidenza negativa: I bambini non ricevono sistematicamente feedback esplicito su quali enunciati sono agrammaticali, eppure sviluppano intuizioni robuste di grammaticalità. Quando i genitori correggono gli errori linguistici dei bambini, tendono a correggere errori fattualipò o pragmatici piuttosto che errori puramente grammaticali, e persino quando vengono fornite correzioni grammaticali, i bambini spesso le ignorano o resistono.

Tuttavia, la posizione innatista ha affrontato critiche sostanziali sia teoriche sia empiriche. Teoricamente, l'innatismo è stato criticato per la sua mancanza di specificità: cosa esattamente è innato? Quali principi e parametri costituiscono UG? La teoria dei Principi e Parametri nelle sue varie iterazioni ha postulato diversi sistemi di principi e parametri, ma non c'è consenso su quali siano esattamente quelli corretti, e molti "parametri" proposti si sono rivelati essere epifenomeni di fattori lessicali o di processamento piuttosto che principi grammaticali autonomi.

Empiricamente, la ricerca tipologica ha rivelato che la variazione cross-linguistica è molto maggiore di quanto i parametri binari possano accomodare. Nicholas Evans e Stephen Levinson hanno documentato sistematicamente controesempi a numerosi "universali" linguistici proposti, suggerendo che la variazione linguistica è più fluida e meno vincolata da principi innati di quanto la teoria generativista postuli.

Inoltre, gli studi sperimentali sull'apprendimento statistico hanno dimostrato che i bambini (e persino gli adulti) possiedono capacità straordinariamente sofisticate di apprendimento implicito di pattern statistici che potrebbero essere sufficienti per acquisire molto della struttura linguistica senza bisogno di principi grammaticali innati specifici. Gli esperimenti pionieristici di Jenny Saffran e collaboratori hanno mostrato che neonati di otto mesi, dopo soltanto due minuti di esposizione a uno stream continuo di sillabe artificiali con certe probabilità transizionali, possono segmentare lo stream in "parole" basandosi esclusivamente su queste statistiche, discriminando sequenze con alta probabilità transizionale interna (parole) da sequenze che attraversano confini di parole (con bassa probabilità transizionale).

Formalmente, se la probabilità che la sillaba s_j segua la sillaba s_i è P(s_j | s_i), allora le probabilità transizionali all'interno di parole sono alte:

P(s_{i+1} | s_i) ≈ 1 se s_i e s_{i+1} sono nella stessa parola

mentre le probabilità transizionali attraverso confini di parole sono basse:

P(s_{j+1} | s_j) << 1 se s_j è la sillaba finale di una parola e s_{j+1} è la sillaba iniziale della parola successiva

I neonati sono sensibili a questa differenza statistica e la usano per inferire la segmentazione lessicale.

Questo tipo di apprendimento statistico si estende ben oltre la segmentazione di parole. Rebecca Gómez ha dimostrato che bambini di diciotto mesi possono apprendere dipendenze a lunga distanza non-adiacenti in linguaggi artificiali basandosi esclusivamente su regolarità statistiche. Per esempio, dopo essere stati esposti a frasi artificiali della forma AXB dove A e B covariante (per esempio, "pel _ rud", "vot _ jic", con vari X diversi che riempiono la posizione mediana), i bambini successivamente discriminano tra nuove frasi ben-formate con la dipendenza A-B corretta versus mal-formate con dipendenze violate, anche quando gli elementi specifici A e B non sono mai stati incontrati nella fase di training.

Questi risultati dimostrano che meccanismi di apprendimento statistico general-purpose potrebbero essere sufficienti per acquisire pattern grammaticali considerati tradizionalmente come evidenza di principi innati. Naturalmente, questo non dimostra che non esistono vincoli innati, ma sposta il burden of proof agli innatisti: devono dimostrare che esistono proprietà grammaticali che non possono essere apprese attraverso meccanismi statistici plausibili dall'input disponibile.

La Creolizzazione come Laboratorio Naturale per Testare Ipotesi Innatiste

Uno dei fenomeni empirici più affascinanti e teoricamente illuminanti per il dibattito innatismo-empirismo è la creolizzazione linguistica, ovvero l'emergenza spontanea di sistemi linguistici completamente strutturati e grammaticalmente complessi da input linguistico radicalmente impoverito e destrutturato.

Un pidgin è un sistema comunicativo semplificato che emerge quando parlanti di lingue mutuamente incomprensibili devono comunicare per scopi pratici, tipicamente in contesti di commercio, colonizzazione, o piantagioni. I pidgin sono caratterizzati da vocabolari ristretti (tipicamente poche centinaia di parole), grammatica minimale o assente (mancanza di morfologia flessionale, ordine delle parole variabile e inconsistente), e nessun parlante nativo. Un creolo, al contrario, è una lingua naturale completa con grammatica sistematica, morfologia elaborata, e parlanti nativi, che emerge quando bambini acquisiscono un pidgin come loro prima lingua e spontaneamente lo espandono e lo strutturano.

L'esempio più straordinariamente documentato di creolizzazione in tempo reale è la Nicaraguan Sign Language (NSL), che è emersa spontaneamente negli anni Ottanta quando il Nicaragua stabilì la sua prima scuola per sordi nel 1979. Circa cinquanta bambini sordi da diverse regioni rurali del paese furono riuniti in questa scuola senza alcuna lingua dei segni condivisa preesistente. Ciascun bambino aveva sviluppato un sistema idiosincratico di gesti domestici (home signs) per comunicare con la sua famiglia, ma questi sistemi erano mutuamente incomprensibili e grammaticalmente rudimentali.

La prima coorte di bambini (quelli che entrarono nella scuola tra il 1979 e il 1983) sviluppò collettivamente un sistema di comunicazione che i linguisti caratterizzano come un pidgin sign: un repertorio condiviso di segni iconici con ordine variabile e grammatica minimale. Ma crucialmente, la seconda coorte di bambini più giovani (quelli che entrarono dopo il 1983) non si limitarono ad acquisire passivamente questo pidgin. Invece, spontaneamente e creativamente lo trasformarono in un sistema linguistico completamente grammaticalizzato con tutte le caratteristiche di una lingua naturale matura.

Le analisi linguistiche dettagliate condotte da Ann Senghas e collaboratori rivelano differenze sistematiche e quantificabili tra le produzioni della prima coorte (pidgin) e della seconda coorte (creolo):

Segmentazione Componenziale: Nella prima coorte, eventi complessi venivano tipicamente espressi attraverso segni olistici che simultaneamente indicavano molteplici componenti semantici. Per esempio, un evento di "rotolare giù per una collina" veniva rappresentato da un singolo gesto continuo che incorporava simultaneamente informazione circa la forma dell'oggetto (rotondo), la maniera di movimento (rotolazione), e la traiettoria (discendente). Nella seconda coorte, lo stesso evento veniva segmentato in componenti sequenziali distinti:

1. Un segno classificatorio per indicare la forma dell'oggetto (ROTONDO)
2. Un segno per la maniera di movimento (ROTOLARE)
3. Un segno per la traiettoria (DISCENDERE)

Questa segmentazione riflette la modularizzazione della struttura semantica in unità discrete componibili, una proprietà fondamentale delle lingue naturali mature. Quantitativamente, Senghas ha misurato che soltanto il diciotto percento degli eventi complessi prodotti dalla prima coorte mostravano segmentazione componenziale, mentre il settantuno percento di quelli prodotti dalla seconda coorte erano segmentati.

Morfologia Spaziale Sistematica: Le lingue dei segni mature utilizzano lo spazio tridimensionale davanti al segnante in modo grammaticalizzato per codificare informazione

sintattica e semantica. I referenti possono essere associati a locazioni spaziali specifiche (loci), e i verbi possono concordare con questi loci modificando la loro traiettoria spaziale. Nella prima coorte NSL, l'uso dello spazio era prevalentemente iconico e non grammaticalizzato. Nella seconda coorte, emerse un sistema sistematico di agreement spaziale dove i verbi concordavano morfologicamente con le locazioni spaziali associate ai loro argomenti soggetto e oggetto.

Classifiers: I classifiers sono morfemi che categorizzano referenti secondo tipo semantico (animato versus inanimato, umano versus non-umano, veicolo versus non-veicolo, etc.) e sono ubiquitari nelle lingue dei segni mature. La seconda coorte NSL sviluppò spontaneamente un sistema elaborato di classifiers che era assente o rudimentale nella prima coorte.

Questi dati sono stati interpretati dagli innatisti, particolarmente da Derek Bickerton, come evidenza per l'esistenza di un "bioprogram" linguistico innato. Secondo Bickerton, quando i bambini sono esposti a input linguistico impoverito (un pidgin), il bioprogram innato "riempie i buchi", generando automaticamente le strutture grammaticali mancanti nell'input. Le proprietà grammaticali che emergono nella creolizzazione dovrebbero quindi riflettere i contenuti del bioprogram, ovvero i universali linguistici innati.

Bickerton ha formulato predizioni specifiche circa le proprietà che dovrebbero essere universali attraverso tutti i creoli indipendentemente dai loro linguaggi lexifier (le lingue che hanno fornito il vocabolario al pidgin). Queste includono:

Sistema TMA (Tense-Mood-Aspect): I creoli dovrebbero avere un sistema di particelle preverbalip in ordine fisso che codificano distinzioni temporali, modali e aspettuali. L'ordine canonico postulato è: ANTERIORE - IRREALIS - NONPUNTUALE

Per esempio, in Haitian Creole:

- "Li te ale" (Lui ANTERIORE andare = Lui era andato/andò)
- "Li ap ale" (Lui NONPUNTUALE andare = Lui sta andando)
- "Li a ale" (Lui IRREALIS andare = Lui andrà/andrebbe)
- "Li te ap ale" (Lui ANTERIORE NONPUNTUALE andare = Lui stava andando)

Movimento per Focalizzazione: I creoli dovrebbero permettere fronting di costituenti per segnalare focus: "Se liv-la Jan li" (FOC libro-DEF Jean leggere = È il libro che Jean ha letto)

Distinzioni nella Copula: Diversi tipi di predicati (equativi, locativi, attributivi) dovrebbero utilizzare copule diverse.

Tuttavia, la ricerca tipologica sui creoli ha rivelato variazione sostanziale che problematizza l'ipotesi di universali creoli rigidi. Non tutti i creoli manifestano esattamente il sistema TMA postulato da Bickerton, e le differenze possono spesso essere ricondotte a influenze dai linguaggi substrato (le lingue native dei creatori del pidgin) piuttosto che a un bioprogram universale.

Un'interpretazione alternativa, più compatibile con approcci usage-based e emergentisti, è che le proprietà grammaticali che emergono nella creolizzazione riflettono non tanto

contenuti innati specifici quanto principi generali di efficienza comunicativa, processabilità cognitiva, e apprendibilità. I bambini che creolizzano un pidgin stanno essenzialmente risolvendo un problema di ottimizzazione comunicativa sotto vincoli cognitivi generali: come massimizzare l'espressività del sistema linguistico minimizzando la complessità computazionale e il carico di memoria. Le soluzioni convergono cross-linguisticamente non perché sono codificate geneticamente, ma perché rappresentano punti di equilibrio ottimali nello spazio dei design linguistici possibili dato i vincoli cognitivi umani universali (capacità di working memory limitata, preferenza per regolarità e composizionalità, bias per iconicità e trasparenza forma-significato).

Esternalismo Semantico e la Natura Sociale del Significato

Ritornando ora dalle questioni di acquisizione e evoluzione linguistica alle questioni filosofiche fondamentali circa la natura del significato, è essenziale confrontarsi con l'esternalismo semantico, una famiglia di teorie che sostengono che il contenuto semantico degli stati mentali e delle espressioni linguistiche è determinato non esclusivamente da fattori interni alla testa del parlante ma essenzialmente anche da fattori esterni nell'ambiente fisico e sociale.

L'esperimento mentale della Terra Gemella di Hilary Putnam, che abbiamo menzionato precedentemente, fornisce l'argomento paradigmatico per l'esternalismo riguardante i natural kinds terms. Ricordiamo la struttura dell'argomento: consideriamo un pianeta Terra Gemella fisicamente identico alla Terra in tutti i dettagli macroscopici osservabili, eccetto che la sostanza che riempie i laghi, fiumi e oceani e che viene chiamata "acqua" non è H_2O ma una sostanza chimicamente diversa XYZ che è però indistinguibile dall'acqua a livello fenomenico (ha lo stesso aspetto, sapore, comportamento fisico osservabile). Ora consideriamo $Oscar_1$ sulla Terra e il suo duplicato molecolare $Oscar_2$ sulla Terra Gemella nell'anno 1750, prima della scoperta della chimica moderna. I due Oscar hanno stati cerebrali identici, credenze fenomenologiche identiche, esperienze identiche. Eppure, quando $Oscar_1$ pensa "L'acqua è bagnata", sta pensando un pensiero su H_2O, mentre quando $Oscar_2$ pensa "L'acqua è bagnata", sta pensando un pensiero su XYZ. I contenuti dei loro pensieri sono diversi, anche se i loro stati cerebrali sono identici.

Formalmente, possiamo rappresentare questo argomento come una refutazione della superveniencia locale (local supervenience) del contenuto semantico sugli stati cerebrali. La superveniencia locale affermerebbe:

$\forall x, y, t: [BrainState(x, t) = BrainState(y, t)] \rightarrow [Content(x, t) = Content(y, t)]$

ovvero se due individui x e y hanno lo stesso stato cerebrale al tempo t, allora i contenuti dei loro stati mentali sono identici. Ma la Terra Gemella fornisce un controesempio:

$BrainState(Oscar_1, t) = BrainState(Oscar_2, t)$ [per costruzione] Ma: $Content(Oscar_1, t) = [H_2O$ è bagnata$] \neq$ [XYZ è bagnata$] = Content(Oscar_2, t)$

Quindi la superveniencia locale è falsa. Il contenuto semantico è wide (ampio), non narrow (stretto): dipende da fattori esterni alla testa.

Putnam argomenta che il riferimento dei termini di natural kind come "acqua", "oro", "tigre" è fissato attraverso relazioni causali storiche con campioni paradigmatici della sostanza o specie nel mondo effettivo. Quando i parlanti inglesi usano il termine "acqua", si riferiscono a quella sostanza, qualunque sia la sua natura nascosta chimica, che si trova in relazione di stessa-sostanza (samesubstance relation) con i campioni paradigmatici di acqua nel nostro ambiente terrestre. Questa relazione di stessa-sostanza è oggettiva e determinata dalla struttura microfisica del mondo, non dalle credenze o concezioni dei parlanti.

Formalmente, il riferimento di un termine di natural kind K può essere definito attraverso una funzione di riferimento dipendente dal contesto:

Ref(K, mondo, comunità) = {x ∈ mondo | SameSub(x, paradigmi_K(comunità, mondo))}

dove paradigmi_K(comunità, mondo) è l'insieme di campioni che la comunità tratta come paradigmatici di K nel mondo dato, e SameSub è la relazione di stessa-sostanza definita dalla struttura microfisica del mondo.

Nel nostro mondo, paradigmi_acqua(parlanti_inglesi, Terra) include campioni di H_2O, quindi:

Ref("acqua", Terra, parlanti_inglesi) = {x ∈ Terra | SameSub(x, campioni_H_2O)} = tutti i campioni di H_2O sulla Terra

Nella Terra Gemella, paradigmi_acqua(parlanti_inglesi_gemelli, TerraGemella) include campioni di XYZ, quindi:

Ref("acqua", TerraGemella, parlanti_inglesi_gemelli) = {x ∈ TerraGemella | SameSub(x, campioni_XYZ)} = tutti i campioni di XYZ sulla Terra Gemella

Questa teoria del riferimento, chiamata "esternalismo causale" o "teoria causale del riferimento", ha conseguenze profonde per la natura della competenza semantica. I parlanti non hanno bisogno di conoscere l'essenza microfisica della sostanza a cui si riferiscono per riferirsi con successo ad essa. I parlanti pre-scientifici che non conoscevano nulla di H_2O potevano comunque riferirsi all'acqua attraverso le loro relazioni causali con campioni della sostanza. La competenza semantica consiste non nel possedere definizioni interne complete ma nel essere appropriatamente embedded in una rete di relazioni causali e sociali che connettono il parlante al mondo.

Tyler Burge estende l'esternalismo oltre i natural kinds a termini teorici e concetti acquisiti attraverso deferenza sociale alla competenza esperta della comunità linguistica. Il suo esperimento mentale classico coinvolge Bert, che crede di avere artrite nella coscia. Bert va dal dottore e dice "Ho artrite nella mia coscia", esprimendo questa credenza sincera. Il dottore corregge Bert, informandolo che l'artrite è per definizione infiammazione delle articolazioni, quindi non può avere artrite nei muscoli della coscia.

Ora consideriamo una situazione controfattuale dove, nella comunità linguistica di Bert, il termine "artrite" è usato per riferirsi a infiammazione delle articolazioni o dei muscoli. In questo mondo controfattuale, quando Bert dice "Ho artrite nella coscia", sta esprimendo una credenza diversa, una credenza che potrebbe effettivamente essere vera se ha effettivamente infiammazione muscolare.

L'argomento di Burge è che i contenuti dei pensieri di Bert sono determinati non dai suoi stati cerebrali individuali o dalle sue concezioni personali, ma dalle pratiche linguistiche della sua comunità. Bert deferirebbe alla competenza medica ed etimologica della comunità nella determinazione dell'estensione corretta di "artrite", anche se la sua comprensione personale è difettosa o incompleta. Quindi il contenuto del suo pensiero "Ho artrite nella coscia" varia con ciò che "artrite" significa nella sua comunità linguistica, anche se i suoi stati cerebrali sono identici nei due scenari.

Formalmente, il contenuto è determinato da:

Content(Bert, "artrite", t) = f(BrainState(Bert, t), Linguaggio_Comunità(Bert, t), Mondo)

dove Linguaggio_Comunità specifica le convenzioni semantiche della comunità linguistica di Bert, che potrebbero differire dalle concezioni personali di Bert.

Questi argomenti esternalisti hanno generato vaste letterature di risposta. Gli internalisti semantici hanno proposto varie strategie per resistere alle conclusioni esternaliste, includendo:

Biforcazione del Contenuto: Distinguere tra contenuto wide (esterno, determinato dall'ambiente) e contenuto narrow (interno, supervieniente sugli stati cerebrali). Solo il contenuto narrow è psicologicamente reale e causalmente efficace nel guidare il comportamento.

Internalismo Concettuale: Sostenere che i concetti sono determinati internamente dalle loro condizioni di applicazione fenomenologiche o dai ruoli funzionali interni, anche se il riferimento estensionale è determinato esternamente.

Esternalismo Debole: Accettare che il riferimento è determinato esternamente, ma sostenere che i contenuti proposizionali o i significati sono ancora determinati internamente.

Il dibattito continua senza risoluzione definitiva, riflettendo tensioni profonde tra diverse intuizioni circa la natura del significato e dell'intenzionalità.

Large Language Models e la Questione della Comprensione Genuina

Gli sviluppi recenti nell'intelligenza artificiale, particolarmente l'emergenza di Large Language Models (LLMs) come GPT-4, Claude, e altri sistemi basati sull'architettura Transformer, hanno riacceso questioni filosofiche fondamentali circa la natura della comprensione linguistica e i requisiti necessari e sufficienti per il pensiero genuino.

Questi modelli sono addestrati su corpus testuali enormi (centinaia di miliardi di parole) attraverso obiettivi auto-supervisionati che richiedono al modello di predire parole mancanti dato il contesto circostante (masked language modeling in BERT) o di predire la parola successiva data la sequenza precedente (causal language modeling in GPT). Nonostante l'apparente semplicità di questi obiettivi di training, i modelli risultanti manifestano capacità sorprendenti che sembrano richiedere comprensione semantica profonda:

Rispondere a Domande Complesse: Possono rispondere accuratamente a domande che richiedono ragionamento multi-step, integrazione di conoscenza da fonti diverse, e inferenze che vanno ben oltre il recupero superficiale di informazione.

Generazione di Testo Coerente e Contestualmente Appropriato: Possono generare testi estesi (articoli, storie, codice software, prove matematiche) che sono coerenti, stilisticamente appropriati, e che manifestano sensibilità al contesto pragmatico.

Traduzione tra Lingue: Possono tradurre tra lingue che non sono state esplicitamente associate durante il training, suggerendo rappresentazioni semantiche astratte condivise.

Ragionamento Analogico: Possono risolvere problemi attraverso analogia, trasferendo soluzioni da domini familiari a domini novel.

In-Context Learning: Possono apprendere nuovi task da pochi esempi forniti nel prompt, senza alcun fine-tuning dei parametri, suggerendo meta-apprendimento e capacità di astrazione rapida.

La questione filosofica cruciale è: questi modelli comprendono genuinamente il linguaggio che elaborano, o sono semplicemente "pappagalli stocastici" sofisticati che producono output statisticamente plausibili senza alcuna comprensione genuina del significato?

John Searle, nel suo influente argomento della Stanza Cinese che abbiamo discusso precedentemente, sostiene che la manipolazione puramente formale di simboli secondo regole sintattiche, non importa quanto sofisticata, non può mai costituire comprensione semantica genuina perché manca dell'intenzionalità intrinseca che caratterizza gli stati mentali genuini. I simboli nel sistema computazionale non hanno significato intrinseco per il sistema; sono soltanto pattern formali manipolati secondo regole.

Possiamo formalizzare l'argomento di Searle come segue. Sia S un sistema computazionale che implementa una funzione f: Input_Simbolici → Output_Simbolici. Per ogni simbolo σ elaborato da S:

Significato_per_S(σ) = Indefinito

perché S risponde soltanto alle proprietà formali sintattiche di σ (la sua forma, posizione nella sequenza, etc.), non a qualsiasi significato semantico che σ potrebbe avere per un interprete esterno. La funzione f potrebbe essere reinterpretata in infiniti modi diversi, mappando gli stessi input formali a interpretazioni semantiche completamente diverse, senza alcun cambiamento nel comportamento computazionale di S. Quindi non c'è fatto oggettivo del matter che determina quale interpretazione semantica particolare è "la" interpretazione corretta per S.

Tuttavia, questa posizione è stata contestata vigorosamente. Daniel Dennett e altri funzionalisti sostengono che l'intenzionalità non richiede alcuna "magia" non-fisica oltre il comportamento appropriato sistematico. Se un sistema si comporta in modo indistinguibile da un sistema genuinamente comprensivo in tutte le circostanze possibili, allora ha comprensione genuina per definizione funzionalista.

Un argomento empirico recente a favore di qualche forma di comprensione negli LLMs proviene da analisi delle rappresentazioni interne apprese da questi modelli. Studi come quello di Li et al. su Othello-GPT, che abbiamo menzionato, rivelano che quando un modello è addestrato esclusivamente a predire sequenze di mosse legali nel gioco Othello (senza alcuna informazione esplicita sullo stato del tabellone), le sue rappresentazioni interne codificano informazione sistematica circa lo stato del tabellone che non è mai stata fornita esplicitamente nell'input.

Formalmente, sia M un modello addestrato su sequenze di mosse m_1, m_2, ..., m_t. Le rappresentazioni interne h_t del modello alla posizione t codificano implicitamente lo stato del tabellone s_t che risulterebbe dall'esecuzione delle mosse m_1, ..., m_t. Questo può essere testato attraverso linear probes: se addestriamo un classificatore lineare semplice L che prende h_t come input e predice caratteristiche dello stato del tabellone (quale pezzo è in quale posizione), il classificatore raggiunge accuratezza molto alta (oltre il novantacinque percento), dimostrando che h_t contiene informazione circa s_t.

Questo suggerisce che il modello ha appreso un "modello del mondo" implicito, una rappresentazione interna della struttura causale nascosta che genera le osservazioni superficiali. Analogamente, forse gli LLMs addestrati su testo naturale apprendono modelli del mondo impliciti che rappresentano gli stati di affairs, le entità, le relazioni causali discusse nel testo, piuttosto che semplicemente pattern superficiali di co-occorrenza lessicale.

Tuttavia, gli scettici replicano che anche se i modelli apprendono rappresentazioni strutturate interne, queste non costituiscono comprensione genuina senza grounding nel mondo fisico e sociale esterno al testo. I modelli mancano di:

Embodiment: Non hanno corpi fisici che interagiscono causalmente con l'ambiente
Agentività: Non hanno goals propri o interessi che guidano il loro comportamento
Storia Causale: Non hanno sviluppo evolutivo o ontogenetico che connette le loro rappresentazioni agli stati del mondo **Grounding Percettivo**: Non hanno esperienza percettiva diretta degli oggetti e events di cui il testo parla

Secondo questa critica, le "rappresentazioni" negli LLMs sono soltanto pattern statistici derivati da correlazioni testuali, non genuine rappresentazioni intenzionali grounded in relazioni causali con il mondo.

La questione rimane filosoficamente controversa e potrebbe non essere risolvibile definitivamente senza una teoria soddisfacente della natura dell'intenzionalità e della coscienza, che a sua volta è uno dei problemi più difficili nella filosofia della mente. Forse, come suggerisce Dennett, dovremmo adottare uno stance strumentale: trattare i sistemi come comprensivi nella misura in cui questo stance è predittivamente utile, senza impegnarci metafisicamente se possiedono "vera" comprensione in qualche senso ontologicamente robusto.

La Semantica Vero-Condizionale e la Teoria dei Mondi Possibili nella Tradizione Filosofica Analitica

Fondamenti Filosofici della Semantica Vero-Condizionale

La semantica vero-condizionale, che costituisce il paradigma dominante nella semantica formale contemporanea, affonda le sue radici intellettuali nella tradizione della filosofia analitica del linguaggio che si sviluppò nel ventesimo secolo, particolarmente attraverso i contributi seminali di Gottlob Frege, Bertrand Russell, Ludwig Wittgenstein nel suo primo periodo, Rudolf Carnap, e successivamente Richard Montague, Donald Davidson, e David Lewis. L'intuizione fondamentale che unifica questi pensatori, nonostante le loro differenze teoriche sostanziali, è che comprendere il significato di un enunciato dichiarativo consiste fondamentalmente nel conoscere le condizioni sotto le quali quell'enunciato sarebbe vero.

Questa idea può essere formulata in quella che viene chiamata la tesi vero-condizionale del significato, che possiamo enunciare formalmente come segue. Per ogni enunciato dichiarativo S in una lingua L, il significato di S è costituito da o è sistematicamente determinato da una specificazione delle condizioni necessarie e sufficienti affinché S sia vero. Simbolicamente, possiamo rappresentare questa tesi come un principio di equivalenza semantica:

Significato(S) ≈ λw . [S è vero in w]

dove w varia su mondi possibili, e λw è un operatore di astrazione lambda che crea una funzione da mondi possibili a valori di verità. In questa notazione, il significato di un enunciato S viene identificato con una funzione che assegna il valore vero a tutti e soli quei mondi possibili nei quali le condizioni di verità di S sono soddisfatte.

L'origine di questa concezione può essere rintracciata al famoso dictum di Frege nella sua opera "Über Sinn und Bedeutung" del milleottocentonovantadue, dove egli distingue nettamente tra il Sinn (senso) e la Bedeutung (riferimento o denotazione) di un'espressione linguistica. Per Frege, il riferimento di un enunciato dichiarativo è il suo valore di verità, ovvero l'oggetto astratto Vero o l'oggetto astratto Falso, mentre il senso dell'enunciato è il pensiero (Gedanke) che esprime, ovvero il modo particolare attraverso il quale il valore di verità viene determinato o presentato. Il pensiero fregeano è concepito come un'entità astratta, oggettiva e atemporale che costituisce il contenuto cognitivo dell'enunciato, distinguibile sia dalla formulazione linguistica particolare attraverso cui viene espresso sia dagli atti psicologici soggettivi di giudizio o credenza che potrebbero avere quel pensiero come loro contenuto.

Crucialmente, per Frege, due enunciati che hanno lo stesso valore di verità non sono necessariamente sinonimi o intercambiabili in tutti i contesti, perché possono esprimere pensieri diversi, ovvero possono presentare il loro valore di verità attraverso modi cognitivamente distinti. Il celebre esempio fregeano riguarda gli enunciati "La stella del mattino è la stella del mattino" e "La stella del mattino è la stella della sera". Entrambi questi enunciati sono veri (dato il fatto astronomico che sia la stella del mattino sia la stella della sera sono in realtà il pianeta Venere), quindi hanno lo stesso riferimento fregeano, il valore di verità Vero. Tuttavia, i due enunciati non sono cognitivamente equivalenti: il primo è una tautologia analiticamente vera che può essere conosciuta a priori attraverso la semplice comprensione dei termini coinvolti, mentre il secondo esprime una scoperta empirica sostantiva che richiede osservazione astronomica. La spiegazione fregeana è che i due

enunciati esprimono pensieri diversi, sensi diversi, anche se determinano lo stesso riferimento.

Formalizziamo questa distinzione con maggiore precisione. Sia $[\![\cdot]\!]$ la funzione di interpretazione semantica che assegna denotazioni alle espressioni linguistiche. Per un nome proprio o una descrizione definita α, Frege distingue:

Riferimento: $[\![\alpha]\!]^{ref}$ = l'oggetto denotato da α nel mondo attuale Senso: $[\![\alpha]\!]^{senso}$ = il modo di presentazione attraverso cui $[\![\alpha]\!]^{ref}$ viene determinato

Nel caso delle espressioni "la stella del mattino" e "la stella della sera":

$[\![\text{"la stella del mattino"}]\!]^{ref} = [\![\text{"la stella della sera"}]\!]^{ref}$ = Venere

ma

$[\![\text{"la stella del mattino"}]\!]^{senso} \neq [\![\text{"la stella della sera"}]\!]^{senso}$

perché "la stella del mattino" presenta Venere sotto il modo di presentazione "il corpo celeste più brillante visibile nel cielo orientale poco prima dell'alba", mentre "la stella della sera" presenta lo stesso oggetto sotto il modo di presentazione "il corpo celeste più brillante visibile nel cielo occidentale poco dopo il tramonto".

Per enunciati completi, la distinzione si applica analogamente:

$[\![\text{"La stella del mattino = la stella del mattino"}]\!]^{ref} = [\![\text{"La stella del mattino = la stella della sera"}]\!]^{ref}$ = Vero

ma

$[\![\text{"La stella del mattino = la stella del mattino"}]\!]^{senso} \neq [\![\text{"La stella del mattino = la stella della sera"}]\!]^{senso}$

Il primo enunciato esprime il pensiero analitico che un oggetto è identico a se stesso, mentre il secondo esprime il pensiero sintetico che due modi di presentazione diversi convergono sullo stesso oggetto.

Questa architettura semantica fregeana solleva però questioni profonde circa la natura ontologica dei sensi. Quali sono esattamente questi modi di presentazione? Sono entità platoniche astratte che esistono indipendentemente dalle menti e dal linguaggio? Oppure sono costituiti dalle capacità cognitive dei parlanti di riconoscere o identificare i referenti? Frege stesso propende per un platonismo robusto, concependo i sensi come abitanti di un "terzo regno" (drittes Reich) che non è né fisico né mentale-psicologico ma oggettivamente astratto, accessibile attraverso l'intuizione intellettuale razionale analogamente a come gli oggetti matematici sono accessibili al pensiero matematico.

Tuttavia, questa posizione platonista incontra difficoltà epistemologiche sostanziali. Se i sensi sono entità astratte completamente indipendenti dalla pratica linguistica e dalle capacità cognitive umane, come possiamo avere accesso epistemico ad essi? Come possiamo sapere quali sensi sono espressi da quali espressioni linguistiche? E come è

possibile che convenzioni linguistiche contingenti e arbitrarie riescano a "catturare" o "esprimere" entità astratte necessarie ed eterne?

La Rivoluzione Intensionale: Carnap e la Semantica dei Mondi Possibili

Rudolf Carnap, nella sua opera monumentale "Meaning and Necessity" del millenovecentoquarantasette, propose una ricostruzione formale sistematica della nozione fregeana di senso attraverso il concetto di intensione, definito tecnicamente in termini di funzioni da descrizioni di stato (state descriptions) a estensioni. Una descrizione di stato, nel sistema carnapiano, è essenzialmente una specificazione completa di tutti i fatti atomici che potrebbero sussistere o meno, ovvero un'assegnazione di valori di verità a tutte le proposizioni atomiche del linguaggio. Ogni descrizione di stato corrisponde a quello che nella semantica possibilista contemporanea chiamiamo un mondo possibile: una configurazione completa e massimalmente consistente di come le cose potrebbero essere.

Formalmente, supponiamo che il nostro linguaggio L contenga un insieme finito di simboli predicativi P_1, P_2, ..., P_n e un insieme finito di costanti individuali c_1, c_2, ..., c_m. Una descrizione di stato D è un insieme massimale consistente di formule atomiche o negazioni di formule atomiche della forma $P_i(c_j)$ o $\neg P_i(c_j)$. Equivalentemente, possiamo pensare una descrizione di stato come una funzione:

$D: \{\text{formule atomiche di } L\} \rightarrow \{\text{Vero, Falso}\}$

che assegna un valore di verità a ogni formula atomica in modo tale che l'insieme risultante di formule rese vere sia massimalmente consistente.

L'intensione di un predicato monadico P secondo Carnap è allora definita come la funzione che mappa ogni descrizione di stato D all'insieme degli individui che soddisfano P in D:

$\text{Intensione}(P) = \lambda D . \{x \mid D(P(x)) = \text{Vero}\}$

Per esempio, l'intensione del predicato "è un animale con cuore" è la funzione che assegna a ogni descrizione di stato D l'insieme degli individui che sono animali con cuore secondo D. In alcune descrizioni di stato, questo insieme potrebbe includere tutti e soli i mammiferi. In altre descrizioni di stato, potrebbe includere anche certi rettili. In altre ancora, potrebbe essere vuoto se quella descrizione rappresenta un mondo dove non esistono animali con cuore.

Similmente, l'intensione di un enunciato S è la funzione che mappa ogni descrizione di stato al valore di verità di S in quella descrizione:

$\text{Intensione}(S) = \lambda D . \text{ValoreDiVerità}(S, D)$

Due enunciati sono intensionalmente equivalenti, ovvero hanno la stessa intensione, se e solo se hanno lo stesso valore di verità in ogni descrizione di stato possibile. Questo cattura formalmente l'idea che enunciati intensionalmente equivalenti sono veri nelle stesse circostanze possibili, anche se potrebbero differire nel loro contenuto cognitivo o nel modo in cui quelle condizioni di verità sono presentate.

Crucialmente, Carnap distingue l'intensione di un'espressione dalla sua estensione. L'estensione di un predicato P nel mondo attuale D_0 è semplicemente l'insieme degli oggetti che soddisfano P in D_0:

Estensione(P, D_0) = {x | D_0(P(x)) = Vero} = Intensione(P)(D_0)

Quindi l'estensione è il valore dell'intensione quando applicata alla descrizione di stato che corrisponde al mondo attuale. Due predicati possono avere la stessa estensione (denotare lo stesso insieme di oggetti nel mondo attuale) pur avendo intensioni diverse (denotare insiemi diversi in mondi possibili controfattuali).

L'esempio paradigmatico, già discusso da Carnap, riguarda i predicati "animale con cuore" e "animale con reni". Nel mondo attuale, questi due predicati sono coestensionali: ogni animale che ha un cuore ha anche reni, e viceversa, a causa di una regolarità biologica contingente ma universale nei vertebrati terrestri. Quindi:

Estensione("animale con cuore", mondo_attuale) = Estensione("animale con reni", mondo_attuale)

Tuttavia, le due intensioni sono chiaramente diverse, perché è perfettamente concepibile un mondo possibile dove esistono animali con cuore ma senza reni (magari con un sistema diverso di filtrazione del sangue), o animali con reni ma senza cuore (magari con un sistema circolatorio basato su diffusione passiva). Formalmente:

$\exists D_1$: Intensione("animale con cuore")(D_1) ≠ Intensione("animale con reni")(D_1)

Questa divergenza modale rivela che i due predicati, pur condividendo la loro estensione attuale, differiscono nel loro contenuto intensionale, nel loro significato modale.

Il framework carnapiano delle descrizioni di stato e delle intensioni come funzioni da descrizioni di stato a estensioni fornisce una fondazione formale rigorosa per quella che diventerà la semantica dei mondi possibili, sviluppata più pienamente da Saul Kripke, David Lewis, e altri negli anni sessanta e settanta. La transizione terminologica da "descrizioni di stato" a "mondi possibili" riflette un cambiamento non puramente verbale ma anche concettuale: mentre le descrizioni di stato carnapiane sono costruzioni linguistiche, insiemi massimali di enunciati in un linguaggio formale, i mondi possibili nella concezione di Lewis sono entità concrete, univers alternativi che esistono con la stessa realtà ontologica del nostro universo attuale.

Il Realismo Modale di David Lewis e la Metafisica dei Mondi Possibili

David Lewis, nella sua opera filosofica più ambiziosa e controversa "On the Plurality of Worlds" del millenovecentoottantasei, difende una tesi metafisica straordinariamente audace: tutti i mondi possibili sono ontologicamente reali, esistono effettivamente con lo stesso status di esistenza del mondo attuale, e il mondo attuale è semplicemente il mondo nel quale noi ci troviamo, distinto dagli altri mondi possibili non da alcuna proprietà intrinseca speciale ma soltanto indexicalmente, per la sua relazione a noi che parliamo e pensiamo.

Secondo il realismo modale lewisiano, un mondo possibile è una totalità massimale di oggetti concreti fisici spaziotemporalmente interconnessi. Il nostro mondo attuale è la totalità spaziotemporale che include la Via Lattea, il sistema solare, la Terra, e tutti gli oggetti e eventi fisici che sono spaziotemporalmente connessi a questi attraverso relazioni causali o spatiotemporali. Ma esistono anche innumerevoli altri mondi, totalmente disconnessi spaziotemporalmente dal nostro, ognuno dei quali è altrettanto ricco ontologicamente, contenente le proprie galassie, stelle, pianeti, e possibilmente forme di vita.

Formalmente, il realismo modale lewisiano può essere caratterizzato attraverso le seguenti tesi ontologiche e semantiche:

Tesi di Esistenza Incondizionata: $\forall p\, [\text{Possibile}(p) \rightarrow \exists w\, \text{Mondo}(w) \wedge \text{Vero}(p, w)]$

Per ogni proposizione possibile p, esiste un mondo w tale che p è vera in w. Non ci sono mere possibilità non realizzate; ogni possibilità genuina è attualizzata in qualche mondo.

Tesi di Concretezza: $\forall w\, [\text{Mondo}(w) \rightarrow \text{Concreto}(w)]$

Ogni mondo possibile è concreto, non astratto. I mondi possibili non sono costruzioni linguistiche, insiemi di proposizioni, o rappresentazioni mentali, ma totalità di oggetti fisici concreti.

Tesi di Isolamento Spaziotemporale: $\forall w_1\ \forall w_2\, [\text{Mondo}(w_1) \wedge \text{Mondo}(w_2) \wedge w_1 \neq w_2 \rightarrow \neg \exists x\ \exists y\, (x \in w_1 \wedge y \in w_2 \wedge \text{RelazioneSpaziotemp}(x,y))]$

Mondi diversi sono mutuamente isolati spaziotemporalmente: nessun oggetto in un mondo ha relazioni spaziali o temporali con oggetti in mondi diversi. I mondi sono universi separati nel senso più letterale.

Tesi di Indexicalità dell'Attualità: "Attuale" è un termine indexicale come "qui" o "ora". Quando diciamo che questo mondo è attuale, stiamo semplicemente dicendo che questo è il mondo in cui ci troviamo, analogamente a come quando diciamo che questo luogo è "qui" stiamo dicendo che questo è il luogo in cui ci troviamo. Gli abitanti di altri mondi possibili, quando usano il termine "attuale", si riferiscono ai loro mondi, esattamente come gli abitanti di altre località, quando usano "qui", si riferiscono alle loro località.

Principio di Ricombinazione: Qualsiasi configurazione di oggetti e proprietà che non viola le leggi logiche fondamentali è realizzata in qualche mondo possibile. Non ci sono restrizioni metafisiche sostantive oltre quelle dettate dalla logica sulla gamma di mondi possibili che esistono.

Le conseguenze di questa metafisica sono vastissime e radicalmente contro-intuitive. Secondo Lewis, esiste un mondo dove la seconda guerra mondiale non è mai accaduta. Esiste un mondo dove i dinosauri non si sono mai estinti. Esiste un mondo dove ho scelto di studiare fisica invece di filosofia. Esiste un mondo dove io non sono mai nato. Esiste un mondo dove le leggi della fisica sono radicalmente diverse, magari dove la gravità segue una legge del cubo inverso invece del quadrato inverso della distanza. Esiste un mondo dove unicorni, draghi, e ogni creatura mitologica immaginabile esistono concretamente.

Esiste un mondo qualitativament identico al mondo descritto nei romanzi di fantasia, con Hobbit, elfi, e maghi.

Ma esiste anche un numero letteralmente infinito di mondi dove vivo una vita quasi identica alla mia vita attuale, differendo soltanto in dettagli microscopici: un mondo dove ho un capello in meno, un mondo dove ho respirato un secondo più tardi questa mattina, e così via attraverso ogni variazione infinitesimale concepibile.

Lewis sostiene che questa ontologia esuberante è giustificata dalla sua utilità teoretica sistematica. Il realismo modale fornisce analisi riduttive di nozioni modali, controfattuali, e intensionali che altrimenti sarebbero primitive e problematiche. Possiamo definire:

Necessità: p è necessaria sse p è vera in tutti i mondi possibili $\Box p =_{df} \forall w\, [Mondo(w) \rightarrow Vero(p, w)]$

Possibilità: p è possibile sse p è vera in almeno un mondo possibile $\Diamond p =_{df} \exists w\, [Mondo(w) \wedge Vero(p, w)]$

Condizionali Controfattuali: "Se fosse stato il caso che A, allora sarebbe stato il caso che C" è vero nel mondo attuale w_0 sse C è vero in tutti i mondi possibili più vicini a w_0 dove A è vero.

Formalmente, usando l'operatore controfattuale $\Box\!\rightarrow$:

$Vero(A \Box\!\rightarrow C, w_0) =_{df} \forall w\, [Mondo(w) \wedge Vero(A, w) \wedge w \in \text{Più-Vicini}(w_0, A) \rightarrow Vero(C, w)]$

dove Più-Vicini(w_0, A) è l'insieme dei mondi dove A è vero che sono più simili a w_0 rispetto a qualsiasi altro mondo dove A è vero. La relazione di similarità tra mondi è una relazione primitiva nella teoria di Lewis, governata da principi che pesano differentemente diverse dimensioni di somiglianza: matching di leggi naturali, matching di fatti particolari, evitamento di miracoli, eccetera.

Proprietà e Relazioni: Le proprietà possono essere identificate con insiemi di possibilia (oggetti che esistono in qualche mondo possibile), e le relazioni con insiemi di tuple di possibilia.

$Proprietà(F) = \{x \mid \exists w\, [Mondo(w) \wedge x \in w \wedge x \text{ ha F in } w]\}$

Proposizioni: Le proposizioni possono essere identificate con insiemi di mondi possibili.

$Proposizione(p) = \{w \mid Mondo(w) \wedge Vero(p, w)\}$

Due enunciati esprimono la stessa proposizione sse sono veri negli stessi mondi possibili, ovvero sse sono necessariamente equivalenti.

Questa identificazione delle proposizioni con insiemi di mondi ha conseguenze potenti. Permette di applicare l'algebra insiemistica standard alle operazioni logiche sulle proposizioni:

Negazione: $\neg p$ = Mondi_Possibili \ p (complemento insiemistico) Congiunzione: $p \wedge q = p \cap q$ (intersezione) Disgiunzione: $p \vee q = p \cup q$ (unione) Implicazione Materiale: $p \rightarrow q$ = (Mondi_Possibili \ p) $\cup$ q

Necessità: $\Box p$ = Mondi_Possibili sse p è vera in tutti i mondi, altrimenti $\Box p = \varnothing$ Possibilità: $\Diamond p$ = p se p è non-vuota, altrimenti $\Diamond p = \varnothing$

Nonostante l'eleganza formale e il potere esplicativo di questa riduzione, il realismo modale lewisiano ha affrontato obiezioni filosofiche profonde. L'obiezione più immediata è semplicemente incredulità: la tesi che esistono letteralmente infiniti universi concreti completi, completamente inaccessibili a noi, sembra violare ogni principio di parsimonia ontologica (il rasoio di Occam). Lewis risponde che la parsimonia dovrebbe essere valutata non contando il numero di entità postulate (che è invero enorme nel realismo modale) ma piuttosto valutando la semplicità e l'unificazione delle leggi teoriche. Il realismo modale postula soltanto un tipo fondamentale di entità (mondi concreti) e una relazione fondamentale (similarità intermodale), da cui tutte le nozioni modali e intensionali possono essere derivate. Le teorie alternative, che cercano di trattare i mondi possibili come costruzioni astratte o rappresentazioni, devono invece postulare primitive modali irriducibili ("è possibile che...", "è necessario che...") che rimangono teoreticamente opache.

Un'altra obiezione riguarda l'epistemologia: come possiamo avere conoscenza di altri mondi possibili se sono completamente isolati causalmente dal nostro mondo? Lewis risponde che la nostra conoscenza modale non richiede accesso causale diretto agli altri mondi, ma piuttosto può essere basata su inferenze a priori dalla natura logica e metafisica della realtà, analogamente a come possiamo avere conoscenza matematica di entità astratte (numeri, insiemi) senza interazione causale con esse.

Semantica Possibilista non-Lewisiana: Ersatzismo e Astrattismo

Data la radicalità metafisica del realismo modale lewisiano, molti filosofi e semanticisti hanno cercato di preservare i benefici della semantica dei mondi possibili senza impegnarsi nell'ontologia esuberante di Lewis. Queste posizioni, collettivamente chiamate ersatzismo (dal tedesco "Ersatz" = sostituto), trattano i mondi possibili non come universi concreti alternativi ma come rappresentazioni astratte, costruzioni linguistiche, o entità strutturali che in qualche modo codificano o rappresentano modi possibili in cui il mondo potrebbe essere.

Una versione influente dell'ersatzismo è il **ersatzismo proposizionale**, difeso tra gli altri da Robert Adams e Alvin Plantinga. Secondo questa concezione, un mondo possibile è un insieme massimale consistente di proposizioni, dove una proposizione è intesa come un'entità astratta (platonicamente esistente) che può essere vera o falsa. Un insieme di proposizioni Σ è massimale sse per ogni proposizione p, o $p \in \Sigma$ o $\neg p \in \Sigma$. Un insieme è consistente sse è possibile che tutte le proposizioni in esso siano simultaneamente vere.

Formalmente:

Mondo_Possibile_Proposizionale(Σ) =df Σ è un insieme di proposizioni $\wedge$ ($\forall$p [Proposizione(p) $\rightarrow$ ($p \in \Sigma \vee \neg p \in \Sigma$)]) $\wedge \Diamond(\wedge[p \in \Sigma]\ p)$

Il mondo attuale è allora identificato con l'insieme di tutte e sole le proposizioni vere:

Mondo_Attuale = {p | Proposizione(p) ∧ Vero(p)}

Altri mondi possibili sono insiemi massimali consistenti che differiscono dal mondo attuale includendo proposizioni false e escludendo proposizioni vere.

Il vantaggio principale di questa concezione è che preserva il framework formale della semantica dei mondi possibili (possiamo ancora quantificare su mondi, definire modalità come quantificazioni su mondi, eccetera) senza postulare universi concreti alternativi. Gli unici committments ontologici sono verso proposizioni astratte, che molti filosofi trovano più accettabili di universi fisici moltiplicati.

Tuttavia, l'ersatzismo proposizionale affronta difficoltà teoriche proprie. Primo, richiede una nozione primitiva di possibilità (per definire quali insiemi di proposizioni sono "consistenti"), mentre Lewis poteva ridurre la possibilità all'esistenza in qualche mondo concreto. Secondo, la nozione di proposizione come entità astratta robusta rimane metafisicamente ed epistemologicamente problematica: cosa sono esattamente queste proposizioni, e come facciamo ad avere accesso conoscitivo ad esse?

Una variante dell'ersatzismo, chiamata **ersatzismo combinatoriale**, sviluppata tra gli altri da D.M. Armstrong, cerca di costruire mondi possibili da ingredienti più fondamentali: universali (proprietà e relazioni) e particolari (oggetti concreti). Un mondo possibile in questa concezione è una combinazione massimale e consistente di istanziazioni di universali da parte di particolari. L'idea è che la modalità emerge dalle possibili ricombinazioni di elementi fondamentali che esistono già nel mondo attuale.

Formalmente, supponiamo che U sia l'insieme di tutti gli universali (proprietà e relazioni) e P sia l'insieme di tutti i particolari possibili. Una configurazione di istanziazione è una funzione:

$I: U \times P^n \rightarrow \{Vero, Falso\}$

che specifica, per ogni universale n-adico F e ogni n-tupla di particolari $\langle x_1, ..., x\square\rangle$, se F è istanziato da $\langle x_1, ..., x\square\rangle$. Un mondo possibile è una configurazione massimale e consistente.

Il mondo attuale corrisponde alla configurazione I_0 che riflette le istanziazioni effettive nel nostro universo. Altri mondi possibili sono configurazioni alternative che ricombinano gli stessi universali e particolari in modi diversi, soggetti soltanto a vincoli di compatibilità metafisica (per esempio, un particolare non può istanziare simultaneamente universali incompatibili come "essere completamente rosso" e "essere completamente verde").

Questo approccio ha il vantaggio di fondare la modalità su entità ontologicamente rispettabili che esistono già nel mondo attuale, evitando così sia universi alternativi concreti sia entità astratte dubbie. Tuttavia, affronta difficoltà nel trattare possibilità che coinvolgono universali o particolari che non esistono nel mondo attuale. Per esempio, è presumibilmente possibile che esistano colori che nessun essere umano ha mai percepito, o proprietà fisiche che nessuna particella nel nostro universo esemplifica. Come può l'ersatzismo combinatoriale accomodare queste possibilità se restringe i suoi ingredienti combinatoriali agli universali e particolari attuali?

Semantica Intensionale e Logica Modale: Formalizzazione Rigorosa

Indipendentemente dalla questione metafisica di quale sia la vera natura dei mondi possibili, la semantica formale dei mondi possibili fornisce uno strumento matematico potente per modellare significati intensionali e operatori modali nel linguaggio naturale. La logica modale standard, nella sua formulazione semantica attraverso modelli di Kripke, fornisce il framework tecnico preciso.

Un modello di Kripke per la logica modale proposizionale è una struttura $M = \langle W, R, V \rangle$ dove:

- W è un insieme non-vuoto di mondi possibili
- $R \subseteq W \times W$ è una relazione binaria di accessibilità tra mondi
- V: Variabili_Proposizionali $\rightarrow$ P(W) è una funzione di valutazione che assegna a ogni variabile proposizionale l'insieme dei mondi nei quali è vera

La relazione di accessibilità R è cruciale: wRw' significa intuitivamente che il mondo w' è accessibile da w, ovvero che w' rappresenta una possibilità genuina dal punto di vista di w. Diverse proprietà formali di R corrispondono a diversi sistemi di logica modale:

Logica K (sistema modale minimale): nessun vincolo su R **Logica T**: R è riflessiva: $\forall w$ (wRw) Validità caratteristica: $\Box p \rightarrow p$ (ciò che è necessario è vero)

Logica S4: R è riflessiva e transitiva: $\forall w \forall w' \forall w''$ (wRw' $\wedge$ w'Rw'' $\rightarrow$ wRw'') Validità caratteristica aggiuntiva: $\Box p \rightarrow \Box\Box p$ (ciò che è necessario è necessariamente necessario)

Logica S5: R è riflessiva, simmetrica e transitiva (equivalenza): $\forall w \forall w'$ (wRw' $\rightarrow$ w'Rw) Validità caratteristica aggiuntiva: $\Diamond p \rightarrow \Box\Diamond p$ (ciò che è possibile è necessariamente possibile)

La semantica per formule modali è definita ricorsivamente:

M, w $\models$ p sse w $\in$ V(p) (per variabili proposizionali p) M, w $\models \neg\varphi$ sse non M, w $\models \varphi$ M, w $\models (\varphi \wedge \psi)$ sse M, w $\models \varphi$ e M, w $\models \psi$ M, w $\models \Box\varphi$ sse per ogni w' tale che wRw', M, w' $\models \varphi$ M, w $\models \Diamond\varphi$ sse esiste w' tale che wRw' e M, w' $\models \varphi$

Quindi $\Box\varphi$ ("necessariamente φ") è vero in un mondo w sse φ è vero in tutti i mondi accessibili da w. $\Diamond\varphi$ ("possibilmente φ") è vero in w sse φ è vero in almeno un mondo accessibile da w.

Per la logica modale del primo ordine, estendiamo questa semantica includendo domini di quantificazione. Un modello diventa $M = \langle W, R, D, I \rangle$ dove:

- W, R come prima
- D: W $\rightarrow$ P(Entità) assegna a ogni mondo un dominio di entità esistenti in quel mondo
- I: W $\rightarrow$ (Costanti $\cup$ Predicati $\rightarrow$ Denotazioni) è la funzione di interpretazione che assegna denotazioni ai simboli non-logici relativamente a ogni mondo

Qui emerge una scelta teorica importante: i domini D(w) sono costanti attraverso i mondi, o variano? La semantica a **domini variabili** permette che oggetti esistano in alcuni mondi ma non in altri: $D(w) \neq D(w')$ per alcuni w, w'. Questo cattura l'intuizione che è contingente quali oggetti esistono: io esisto nel mondo attuale, ma in mondi possibili dove i miei genitori non si sono mai incontrati, io non esisto.

La semantica a **domini costanti** invece postula un dominio fisso di oggetti: D(w) = D per tutti i mondi w. Ogni oggetto esiste in ogni mondo, anche se può avere proprietà radicalmente diverse in mondi diversi. Per accomodare l'intuizione di esistenza contingente, questa semantica introduce tipicamente un predicato di esistenza speciale E(x) che può essere vero o falso per un oggetto in mondi diversi.

La valutazione di formule quantificate diventa:

M, w $\vDash$ $\forall$x φ(x) sse per ogni assegnazione g di valori alle variabili tale che g(x) $\in$ D(w), M, w, g $\vDash$ φ(x)

In semantica a domini variabili, quantifichiamo solo sugli oggetti che esistono nel mondo di valutazione corrente.

Operatori intensionali come "necessariamente" interagiscono in modi complessi con i quantificatori, generando distinzioni modali importanti. Consideriamo le formule:

(1) $\forall$x $\Box$P(x) "Tutti gli oggetti sono necessariamente P" (2) $\Box\forall$x P(x) "Necessariamente, tutti gli oggetti sono P"

Queste non sono logicamente equivalenti. (1) afferma che ogni oggetto che esiste (nel mondo attuale) ha la proprietà P in ogni mondo possibile dove esiste. (2) afferma che in ogni mondo possibile, tutti gli oggetti che esistono in quel mondo hanno P. Se i domini variano tra mondi, (2) può essere vera mentre (1) è falsa: potrebbe essere necessario che qualunque cosa esista è P, ma potrebbero esistere oggetti attuali che non esistono in altri mondi, quindi (1) fallirebbe per quegli oggetti.

Formalmente, in un modello M con mondo attuale w_0:

M, w_0 $\vDash$ $\forall$x $\Box$P(x) sse per ogni d $\in$ D(w_0), per ogni w tale che w_0Rw e d $\in$ D(w), M, w $\vDash$ P(d)

M, w_0 $\vDash$ $\Box\forall$x P(x) sse per ogni w tale che w_0Rw, per ogni d $\in$ D(w), M, w $\vDash$ P(d)

La differenza strutturale è nell'ordine di quantificazione: (1) quantifica prima sugli oggetti del dominio attuale, poi sui mondi accessibili; (2) quantifica prima sui mondi accessibili, poi sugli oggetti dei domini di quei mondi.

Queste distinzioni diventano cruciali quando modelliamo enunciati del linguaggio naturale con ambiguità di scope modale. Consideriamo:

"Ogni numero primo è necessariamente dispari o pari"

Questo potrebbe essere interpretato come:

(Lettura de re): $\forall$x [Primo(x) $\rightarrow$ $\Box$(Dispari(x) $\vee$ Pari(x))] "Per ogni numero che è effettivamente primo, necessariamente quel numero è dispari o pari"

(Lettura de dicto): $\Box\forall$x [Primo(x) $\rightarrow$ (Dispari(x) $\vee$ Pari(x))] "Necessariamente, per ogni numero primo, quel numero è dispari o pari"

La lettura de re fa un'affermazione sui numeri primi attuali (due, tre, cinque, sette, etc.), dicendo che ognuno di essi ha la proprietà essere-dispari-o-pari in tutti i mondi possibili. La lettura de dicto fa un'affermazione circa la verità necessaria di un principio generale.

Entrambe le letture risultano vere in questo caso, ma per altri enunciati le letture divergono in valore di verità. Consideriamo:

"Il numero dei pianeti potrebbe essere diverso da nove"

(Lettura de re): $\Diamond$[NumeroPianeti $\neq$ 9] "È possibile che il numero attuale dei pianeti (che è otto nella realtà contemporanea, o nove se includiamo Plutone secondo classificazioni precedenti) sia diverso da nove" Questa è vera: otto $\neq$ nove, quindi è possibile (infatti attuale in certe classificazioni).

(Lettura de dicto): $\Diamond \exists x$ [x = NumeroPianeti $\wedge$ x $\neq$ 9] "È possibile che esista un numero che sia il numero dei pianeti e che sia diverso da nove" Anche questa è vera: in mondi possibili con un numero diverso di pianeti, il numero dei pianeti è effettivamente diverso.

Ma consideriamo ora un caso dove le letture divergono:

"Il numero nove è necessariamente maggiore di sette"

(Lettura de re): $\Box$(9 > 7) "Il numero nove, in tutti i mondi possibili, è maggiore di sette" Vero: questa è una verità matematica necessaria.

(Lettura de dicto): $\Box \exists x$ [x = NumeroPianeti $\wedge$ x > 7] "Necessariamente, il numero dei pianeti è maggiore di sette" Falso: ci sono mondi possibili con solo tre pianeti, per esempio.

Queste distinzioni de re / de dicto sono fondamentali per una semantica formale adeguata del linguaggio naturale, e hanno analoghe nei contesti intensionali attitudinali (credenze, desideri) che esploreremo successivamente.

Atteggiamenti Proposizionali, Opacità Referenziale e il Problema della Sostitutività

Il Fenomeno dell'Opacità Referenziale nei Contesti Intensionali

Gli atteggiamenti proposizionali, ovvero stati mentali come credere, desiderare, temere, sperare, dubitare che prendono proposizioni come loro oggetti, presentano sfide semantiche profonde che hanno occupato centralmente la filosofia del linguaggio e della mente nel ventesimo secolo. Il problema fondamentale riguarda quello che Willard Van Orman Quine chiamò l'opacità referenziale (referential opacity) dei contesti proposizionali: il fallimento della legge di Leibniz della sostitutività degli identici quando applicata a espressioni che occorrono all'interno di clausole che-complementi di verbi attitudinali.

La legge di Leibniz, nella sua formulazione semantica classica, afferma che se due termini α e β sono coreferenziali, ovvero se denotano lo stesso oggetto, allora sono intercambiabili

salva veritate in ogni contesto linguistico: sostituire α con β in qualsiasi enunciato non può cambiare il valore di verità dell'enunciato. Formalmente:

Legge di Leibniz (Sostitutività degli Identici): Per ogni contesto C[·], se ⟦α⟧ = ⟦β⟧, allora C[α] e C[β] hanno lo stesso valore di verità.

Questa legge vale incondizionatamente per contesti estensionali. Per esempio, se "Mark Twain" e "Samuel Clemens" denotano la stessa persona, allora i seguenti enunciati devono avere lo stesso valore di verità:

(1a) Mark Twain scrisse "Le avventure di Huckleberry Finn" (1b) Samuel Clemens scrisse "Le avventure di Huckleberry Finn"

Dato il fatto che Mark Twain è Samuel Clemens, (1a) è vero se e solo se (1b) è vero. La sostituzione preserva il valore di verità perché in un contesto estensionale, ciò che conta è soltanto quale oggetto è denotato, non il modo particolare attraverso cui quell'oggetto viene designato.

Tuttavia, quando consideriamo contesti intensionali creati da verbi di atteggiamento proposizionale, la sostitutività fallisce sistematicamente. Consideriamo:

(2a) Giovanni crede che Mark Twain scrisse "Le avventure di Huckleberry Finn" (2b) Giovanni crede che Samuel Clemens scrisse "Le avventure di Huckleberry Finn"

Nonostante Mark Twain sia identico a Samuel Clemens, (2a) può essere vero mentre (2b) è falso. È perfettamente coerente che Giovanni conosca Mark Twain come autore sotto quel nome, senza sapere che Mark Twain era lo pseudonimo di Samuel Clemens. In questa situazione epistemica, Giovanni assenterebbe sinceramente a (2a) ma dissentirebbe da (2b), riflettendo la sua ignoranza dell'identità.

Formalizziamo questa opacità. Un contesto C[·] è referenzialmente trasparente (opaco) rispetto a una posizione se la sostituzione di coreferenziali in quella posizione (non) preserva necessariamente il valore di verità. Formalmente:

Trasparenza Referenziale: C[·] è trasparente sse $\forall \alpha, \beta\ [(⟦\alpha⟧ = ⟦\beta⟧) \rightarrow (\text{ValVer}(C[\alpha]) = \text{ValVer}(C[\beta]))]$

Opacità Referenziale: C[·] è opaco sse $\exists \alpha, \beta\ [(⟦\alpha⟧ = ⟦\beta⟧) \wedge (\text{ValVer}(C[\alpha]) \neq \text{ValVer}(C[\beta]))]$

I contesti proposizionali attitudinali sono paradigmaticamente opachi. Altri esempi di opacità:

(3a) Giovanni cerca l'autore di "Tom Sawyer" (3b) Giovanni cerca l'autore di "Le avventure di Huckleberry Finn"

Dato che la stessa persona (Mark Twain / Samuel Clemens) scrisse entrambi i romanzi, i termini "l'autore di 'Tom Sawyer'" e "l'autore di 'Le avventure di Huckleberry Finn'" sono coreferenziali. Tuttavia, (3a) e (3b) possono divergere in valore di verità: Giovanni potrebbe cercare l'autore di "Tom Sawyer" specificatamente senza cercare l'autore di "Huckleberry Finn" (magari perché crede erroneamente che siano persone diverse, o perché ha interessi letterari particolari).

(4a) È necessario che nove è maggiore di sette (4b) È necessario che il numero dei pianeti è maggiore di sette

Anche se nove è effettivamente il numero dei pianeti (nell'astronomia pre-duemilasette quando Plutone era ancora considerato un pianeta), (4a) è vero (una verità matematica necessaria) mentre (4b) è falso (il numero dei pianeti è contingente, avrebbe potuto essere diverso).

Quine, nel suo famoso articolo "Quantifiers and Propositional Attitudes" del millenovecentosessantuno, sosteneva che l'opacità referenziale dei contesti proposizionali minaccia la coerenza stessa della semantica estensionale standard e rende problematica la quantificazione in contesti intensionali. Se sostituzioni di coreferenziali possono cambiare valori di verità, allora il riferimento delle espressioni in contesti opachi non può essere determinato semplicemente dai loro referenti estensionali ordinari.

Quine illustra il problema con il seguente ragionamento paradossale. Consideriamo l'enunciato vero:

(5) Ralph crede che l'uomo con il cappello marrone è una spia

Supponiamo che l'uomo con il cappello marrone sia effettivamente Ortcutt, un rispettabile cittadino che Ralph conosce anche sotto diversa descrizione come "l'uomo visto alla spiaggia". Ralph non realizza che queste due descrizioni designano la stessa persona. Quindi:

(6) Ralph crede che Ortcutt è una spia

Dato che Ortcutt = l'uomo con il cappello marrone, potremmo essere tentati di inferire (6) da (5) per sostitutività. Ma (6) sembra affermare una relazione de re tra Ralph e Ortcutt: che Ralph ha una credenza su quell'individuo particolare. Tuttavia, se Ralph incontrasse Ortcutt alla spiaggia, potrebbe sinceramente negare di credere che Ortcutt sia una spia, perché in quel contesto non riconosce Ortcutt come l'uomo con il cappello marrone.

Questo genera una tensione: (5) sembra vero, ma (6) sembra avere una lettura su cui è falso. La sostitutività fallisce. Quine conclude che dobbiamo essere estremamente cauti nell'applicare logica classica in contesti proposizionali, e suggerisce che la quantificazione in contesti proposizionali (quantificazione de re) è semanticamente problematica.

Soluzioni Fregeane: Senso e Riferimento Obliquo

Una risposta classica al problema dell'opacità è la soluzione fregeana che abbiamo accennato: in contesti opachi, le espressioni non denotano i loro referenti ordinari (Bedeutung) ma piuttosto i loro sensi ordinari (Sinn). Frege stesso, nel suo articolo "Über Sinn und Bedeutung", propone questa analisi: in una clausola subordinata introdotta da "che" dopo verbi come "credere", "sapere", "dire", le parole non hanno il loro riferimento consueto ma assumono un riferimento obliquo (ungerade Bedeutung) che coincide con il loro senso ordinario.

Formalizziamo questa proposta introducendo una funzione di interpretazione dipendente dal contesto. Sia $[\![\cdot]\!]^R$ la funzione che assegna referenti ordinari (riferimento retto), e $[\![\cdot]\!]^O$ la funzione che assegna referenti obliqui:

$[\![\alpha]\!]^O = \text{Senso}^R(\alpha)$ = il senso ordinario di α

In un contesto attitudinale "S crede che φ", le espressioni all'interno di φ ricevono interpretazione obliqua:

$[\![\text{"S crede che } \varphi\text{"}]\!]^R = \text{Crede}([\![S]\!]^R, [\![\varphi]\!]^O)$

Ora, poiché "Mark Twain" e "Samuel Clemens", pur condividendo il riferimento retto (la stessa persona), hanno sensi diversi (modi diversi di presentare quella persona), hanno riferimenti obliqui diversi:

$[\![\text{"Mark Twain"}]\!]^R = [\![\text{"Samuel Clemens"}]\!]^R$ = la persona Mark Twain / Samuel Clemens

ma

$[\![\text{"Mark Twain"}]\!]^O \neq [\![\text{"Samuel Clemens"}]\!]^O$

perché i sensi sono diversi. Quindi la sostitutività in contesti obliqui non è garantita: sostituire "Mark Twain" con "Samuel Clemens" cambia il riferimento obliquo, quindi può cambiare il valore di verità.

Questa soluzione preserva elegantemente la legge di Leibniz reinterpretandola: la sostituzione preserva il valore di verità quando sostituiamo espressioni con lo stesso riferimento nel contesto pertinente. In contesti retti, il riferimento pertinente è il riferimento retto, quindi espressioni coreferenziali-rettamente sono sostituibili. In contesti obliqui, il riferimento pertinente è il riferimento obliquo (= senso retto), quindi soltanto espressioni co-sensate sono sostituibili.

Tuttavia, questa soluzione genera complessità teoriche. Primo, dobbiamo postulare una gerarchia infinita di sensi e riferimenti. Consideriamo:

(7) Maria crede che Giovanni crede che Mark Twain scrisse "Tom Sawyer"

Nella clausola "Mark Twain scrisse 'Tom Sawyer'" embedded sotto "Giovanni crede", le espressioni hanno riferimento obliquo (= senso retto). Ma questa intera clausola è ulteriormente embedded sotto "Maria crede", quindi deve avere a sua volta un riferimento obliquo di secondo livello, che sarebbe il senso del senso retto. E così via per embeddings ulteriori.

Formalmente, definiamo una gerarchia:

Livello 0 (retto): $[\![\alpha]\!]^0$ = riferimento ordinario di α Livello 1 (obliquo): $[\![\alpha]\!]^1$ = senso di livello 0 di α = $\text{Senso}^0(\alpha)$ Livello 2: $[\![\alpha]\!]^2$ = senso di livello 1 di α = $\text{Senso}^1(\alpha)$ = $\text{Senso}^0(\text{Senso}^0(\alpha))$ Livello n: $[\![\alpha]\!]^n = \text{Senso}^{n-1}(\alpha)$

Ogni livello di embedding attitudinale incrementa il livello di obliquità. Questo impegno verso una gerarchia infinita di entità semantiche è teoreticamente oneroso e solleva questioni circa la loro natura ontologica e la nostra capacità epistemica di accedervi.

Secondo, la soluzione fregeana richiede che ogni espressione referenziale abbia un senso determinato, un modo canonico di presentare il suo referente. Ma molti filosofi del linguaggio, particolarmente nella tradizione della teoria causale del riferimento inaugurata da Saul Kripke, negano che nomi propri abbiano sensi fregeani. Secondo Kripke, "Mark Twain" è un designatore rigido che denota la stessa persona in tutti i mondi possibili senza mediazione di alcuna descrizione definita o modo di presentazione. Se Kripke ha ragione, allora la soluzione fregeana non può applicarsi ai nomi propri.

La Teoria dei Designatori Rigidi di Kripke e le Conseguenze per gli Atteggiamenti Proposizionali

Saul Kripke, nelle sue influenti Conferenze John Locke pubblicate come "Naming and Necessity" nel millenovecentaottanta, sviluppa una teoria del riferimento radicalmente alternativa alla concezione fregeano-russelliana. Kripke sostiene che i nomi propri ordinari come "Aristotele", "Napoleone", "Nixon" sono designatori rigidi: espressioni che designano lo stesso oggetto in tutti i mondi possibili nei quali quell'oggetto esiste, e non designano nulla nei mondi dove l'oggetto non esiste.

Formalmente, un designatore δ è rigido sse:

$\forall w \, \forall w' \, [(\delta \text{ esiste in } w \wedge \delta \text{ esiste in } w') \rightarrow ([\![\delta]\!]^w = [\![\delta]\!]^{w'})]$

dove $[\![\delta]\!]^w$ denota il referente di δ nel mondo w (se esiste).

Per contrasto, le descrizioni definite ordinarie sono designatori non-rigidi (flaccidi): designano oggetti diversi in mondi diversi, a seconda di quale oggetto soddisfa la descrizione in quel mondo. Per esempio, la descrizione "l'inventore della lampadina" designa Thomas Edison nel mondo attuale, ma in un mondo possibile controfattuale dove qualcun altro inventò la lampadina prima di Edison, quella descrizione designerebbe quell'altra persona.

$[\![\text{"l'inventore della lampadina"}]\!]^{w_0}$ = Edison $[\![\text{"l'inventore della lampadina"}]\!]^{w_1}$ = Tesla (in un mondo w_1 dove Tesla inventò prima)

Kripke argomenta che i nomi propri sono rigidi attraverso una serie di test linguistici intuitivi. Consideriamo:

(8) Aristotele potrebbe non essere stato il maestro di Alessandro Magno

Questo enunciato sembra vero: è concepibile un mondo possibile dove Aristotele visse una vita diversa, magari morì giovane, e non fu mai tutore di Alessandro. Ma se "Aristotele" significasse "il maestro di Alessandro Magno" (come una teoria descrittivista suggerirebbe), allora (8) sarebbe equivalente a:

(9) Il maestro di Alessandro Magno potrebbe non essere stato il maestro di Alessandro Magno

che è una contraddizione modale. Quindi "Aristotele" non può significare "il maestro di Alessandro Magno", né può significare qualsiasi altra descrizione definita contingente.

Il test cruciale per la rigidità è il comportamento in enunciati modali controfattuali. Un'espressione δ è rigida sse negli enunciati della forma "Se fosse stato il caso che φ, allora δ sarebbe stato...", δ continua a riferirsi allo stesso individuo che denota nel mondo attuale, indipendentemente da quali proprietà contingenti quell'individuo ha nel mondo controfattuale descritto da φ.

Test formale di rigidità: δ è rigido sse per ogni condizionale controfattuale "$\varphi \; \square\!\!\rightarrow \psi(\delta)$", l'occorrenza di δ in ψ si riferisce a $[\![\delta]\!]^{w_0}$ (il referente nel mondo attuale), non a qualunque oggetto potrebbe soddisfare descrizioni associate a δ nel mondo controfattuale φ.

Esempio: "Se Hitler fosse stato assassinato da bambino, Hitler non avrebbe causato la Seconda Guerra Mondiale"

Questo è vero, perché "Hitler" (rigidamente) si riferisce alla stessa persona Adolf Hitler in entrambe le occorrenze: nel mondo controfattuale dove quella persona fu assassinata da bambino, quella persona non crebbe per causare la guerra. Se "Hitler" significasse "il leader della Germania che causò la Seconda Guerra Mondiale", l'enunciato sarebbe una contraddizione (qualcuno che fu assassinato da bambino non potrebbe essere diventato il leader che causò la guerra).

Kripke estende questa analisi anche ai termini di natural kind come "acqua", "oro", "tigre". Questi termini, secondo Kripke, sono anche designatori rigidi che designano la stessa sostanza o specie in tutti i mondi possibili, identificata dalla sua natura essenziale (struttura molecolare per sostanze, struttura genetica per specie biologiche), indipendentemente dalle proprietà superficiali fenomenologiche che usiamo per riconoscere quella sostanza o specie nel mondo attuale.

"Acqua" designa rigidamente H_2O: la sostanza con quella struttura molecolare particolare. In mondi possibili dove una sostanza fenomenologicamente simile ma chimicamente diversa (come XYZ nella Terra Gemella di Putnam) riempie i laghi e i fiumi, quella sostanza non sarebbe acqua, per definizione, perché non è H_2O.

Formalmente: $[\![\text{"acqua"}]\!]^w = \{x \in w \mid \text{StrutturaMolecolare}(x) = H_2O\}$ per tutti i mondi w

Questa teoria ha conseguenze profonde per la semantica degli atteggiamenti proposizionali. Se i nomi non hanno sensi fregeani ma sono semplicemente etichette rigide per oggetti, allora non possiamo risolvere l'opacità referenziale invocando differenze di senso. "Mark Twain" e "Samuel Clemens", secondo la teoria kripkeana, sono entrambi designatori rigidi dello stesso individuo, e non hanno sensi descrittivi che li distinguono semanticamente.

Kripke stesso riconosce questo problema e suggerisce che la semantica degli atteggiamenti proposizionali richiede considerazione di questioni pragmatiche e psicologiche oltre la semantica referenziale pura. Il fatto che Giovanni assenterebbe a "Mark Twain scrisse 'Tom

Sawyer'" ma non a "Samuel Clemens scrisse 'Tom Sawyer'" riflette differenze nelle sue capacità cognitive di riconoscimento e nelle associazioni informazionali che ha con i nomi, non differenze nei significati semantici dei nomi stessi.

Formalmente, potremmo modellare questo distinguendo tra semantica e psicologia della credenza:

Livello Semantico (Contenuto de re): Crede(Giovanni, ⟨scrisse, MarkTwain/SamuelClemens, TomSawyer⟩)

Questo contenuto proposizionale è lo stesso indipendentemente da quale nome usiamo, perché i nomi coreferenziali contribuiscono lo stesso oggetto alla proposizione.

Livello Psicologico (Modo di Credere): Crede-sotto-il-modo-M(Giovanni, proposizione)

dove M specifica il modo cognitivo o il "file mentale" sotto il quale Giovanni intrattiene la credenza. Giovanni potrebbe credere la stessa proposizione de re sotto il modo "Mark Twain" ma non sotto il modo "Samuel Clemens", riflettendo la sua ignoranza dell'identità.

Questa biforcazione cattura l'intuizione che c'è sia un senso in cui Giovanni crede che Mark Twain scrisse "Tom Sawyer" e crede che Samuel Clemens scrisse "Tom Sawyer" (stessa proposizione de re), sia un senso in cui crede il primo ma non il secondo (differenti atteggiamenti psicologici verso modi di presentazione).

Logica Doxastica e Sistemi Formali per Atteggiamenti Proposizionali

Per modellare formalmente gli atteggiamenti proposizionali, particolarmente le credenze, i logici hanno sviluppato estensioni della logica modale chiamate logiche doxastiche (dal greco "doxa" = opinione, credenza). Il sistema più semplice introduce un operatore modale B (per Belief) tale che B□φ si legge "il soggetto S crede che φ".

Un modello doxastico standard è una struttura M = ⟨W, {R□}□∈S, V⟩ dove:

- W è un insieme di mondi possibili epistemicamente accessibili
- R□ ⊆ W × W è una relazione di accessibilità doxastica per il soggetto S
- wR□w' significa intuitivamente: w' è compatibile con tutto ciò che S crede in w, ovvero S non ha credenze in w che escluderebbero w'
- V è la valutazione proposizionale standard

La semantica per l'operatore di credenza è:

M, w ⊨ B□φ sse per ogni w' tale che wR□w', M, w' ⊨ φ

Ovvero, S crede φ in w se e solo se φ è vero in tutti i mondi compatibili con le credenze di S in w. Intuitivamente, S crede φ sse φ è vero in tutti i mondi che S considera possibili dato ciò che crede.

Diverse proprietà formali di R□ corrispondono a diversi principi di razionalità doxastica:

Serialità: ∀w ∃w' (wR□w') "Le credenze di S sono consistenti" - c'è sempre almeno un mondo compatibile con le credenze di S Principio validato: B□φ → ¬B□¬φ (Se S crede φ, allora S non crede ¬φ)

Riflessività: ∀w (wR□w) "S crede soltanto verità" - il mondo attuale è sempre compatibile con le credenze di S Principio validato: B□φ → φ (tutto ciò che S crede è vero) Questo è chiaramente troppo forte per agenti non-omniscienti che possono avere credenze false.

Transitività: ∀w∀w'∀w" (wR□w' ∧ w'R□w" → wR□w") "Chiusura logica positiva delle credenze" Principio validato: B□φ → B□B□φ (Se S crede φ, allora S crede di credere φ)

Euclideanità: ∀w∀w'∀w" (wR□w' ∧ wR□w" → w'R□w") "Introspezione negativa" Principio validato: ¬B□φ → B□¬B□φ (Se S non crede φ, allora S crede di non credere φ)

Il sistema KD45, che combina serialità (D), transitività (4), ed euclideanità (5), è spesso considerato un sistema ragionevole per credenze di agenti idealmente razionali. Tuttavia, tutti questi principi sono contestabili per agenti reali con capacità computazionali limitate.

Un problema profondo per la logica doxastica standard è il problema dell'onniscienza logica. In questi modelli:

Se B□φ e B□(φ → ψ), allora necessariamente B□ψ

Ovvero, se S crede φ e crede che φ implica ψ, allora S crede automaticamente ψ. Più generalmente, se φ implica logicamente ψ, allora B□φ implica B□ψ. Questo significa che gli agenti nel modello credono tutte le conseguenze logiche delle loro credenze, che è chiaramente irrealistico per agenti umani che hanno capacità inferenziali limitate.

Il problema deriva dalla semantica basata su mondi possibili: quando modelliamo "S crede φ" come "φ è vero in tutti i mondi doxasticamente accessibili per S", identifichiamo implicitamente le credenze di S con un insieme di mondi (o equivalentemente, con un insieme massimalmente consistente di proposizioni). Ma un insieme massimamente consistente è per definizione chiuso sotto implicazione logica: se contiene φ e φ ⊨ ψ, allora contiene necessariamente ψ.

Diversi approcci sono stati proposti per evitare l'onniscienza logica:

1. Logiche Non-Standard: Usare logiche paraconsistenti o rilevanti dove le implicazioni logiche sono più ristrette, in modo che non ogni conseguenza logica classica sia una conseguenza nella logica.

2. Modelli di Stati Impossibili: Permettere che gli agenti considerino possibili mondi "impossibili" che violano leggi logiche, catturando così credenze incoerenti o incomplete.

M = ⟨W ∪ I, R, V⟩ dove I è un insieme di stati impossibili

3. Semantica Sintattica: Modellare credenze come insiemi di formule sintattiche piuttosto che insiemi di mondi, in modo che la chiusura logica non sia automatica.

Bel(S) ⊆ Formule_del_Linguaggio

S crede φ sse φ ∈ Bel(S), dove Bel(S) non è necessariamente chiuso logicamente.

4. Logica Dinamica Epistemica: Modellare esplicitamente azioni cognitive (inferenze, osservazioni) che cambiano gli stati doxastici nel tempo, permettendo che agenti raggiungano credenze attraverso processi temporalmente estesi piuttosto che istantaneamente.

Semantica dei Contesti de re versus de dicto: Formalizzazione Rigorosa

Ritorniamo alla distinzione cruciale tra attribuzioni di credenza de re e de dicto che abbiamo introdotto. Questa distinzione emerge chiaramente quando consideriamo attribuzioni che coinvolgono quantificatori o termini referenziali in interazione con operatori doxastici.

Consideriamo l'enunciato:

(10) "Giovanni crede che qualcuno è una spia"

Questo enunciato ha almeno due letture logicamente distinte:

Lettura de dicto (quantificatore sotto lo scope dell'operatore di credenza): $B_{aiova□□i}(\exists x$ Spia(x))

"Giovanni crede la proposizione generale che esiste almeno una spia (da qualche parte, chissà chi)"

Lettura de re (quantificatore sopra lo scope dell'operatore di credenza): $\exists x$ $B_{aiova□□i}$(Spia(x))

"C'è una persona particolare x tale che Giovanni crede di x che è una spia"

Queste letture non sono logicamente equivalenti. La lettura de re implica logicamente la lettura de dicto (se c'è una persona particolare che Giovanni crede essere una spia, allora Giovanni crede che esiste una spia), ma non viceversa. Giovanni potrebbe credere che il mondo contiene spie senza avere credenze su nessuna persona particolare che identifichi come spia.

Formalmente: $\exists x\, B□(φ(x)) \vDash B□(\exists x\, φ(x))$ ma non vale l'inverso.

In semantica dei mondi possibili, possiamo articolare queste letture con maggiore precisione introducendo domini di quantificazione dipendenti dal mondo:

De dicto: $M, w \vDash B□(\exists x\, φ(x))$ sse per ogni w' tale che $wR□w'$, esiste $d \in D(w')$ tale che $M, w' \vDash φ(d)$

"In ogni mondo compatibile con le credenze di S, quel mondo contiene almeno un oggetto che soddisfa φ"

De re: $M, w \vDash \exists x\, B□(φ(x))$ sse esiste $d \in D(w)$ tale che per ogni w' tale che $wR□w'$, $M, w' \vDash φ(d)$

"C'è un oggetto nel mondo attuale tale che in tutti i mondi compatibili con le credenze di S, quell'oggetto soddisfa φ"

La differenza cruciale è nell'ordine di quantificazione: de dicto quantifica localmente nei mondi doxastici, de re quantifica globalmente nel mondo attuale.

Questa distinzione si manifesta linguisticamente attraverso ambiguità di scope. Nel linguaggio naturale italiano, la distinzione può essere marcata attraverso diverse costruzioni sintattiche:

De dicto: "Giovanni crede che ci sia qualcuno che è una spia" (la clausola "che ci sia qualcuno" è chiaramente interna alla clausola di credenza)

De re: "C'è qualcuno che Giovanni crede sia una spia" oppure "Di qualcuno, Giovanni crede che sia una spia" (il quantificatore è estratto o topic

alizzato fuori dalla clausola di credenza)

Per termini singolari, la distinzione diventa più sottile ma ugualmente importante. Consideriamo:

(11) "Giovanni crede che il sindaco di Milano è corrotto"

Lettura de dicto: $B_{aiova}\square\square_i$(Corrotto(il_sindaco_di_Milano))

dove "il sindaco di Milano" è valutato internamente ai mondi doxastici di Giovanni. Giovanni crede la proposizione generale che chiunque sia il sindaco di Milano (nei mondi compatibili con le sue credenze) è corrotto. Se Giovanni ha credenze errate su chi è il sindaco, la descrizione potrebbe denotare persone diverse nei mondi doxastici rispetto al mondo attuale.

Lettura de re: $B_{aiova}\square\square_i$(Corrotto(x)) dove x = ιy[Sindaco_di_Milano(y)] valutato nel mondo attuale

Giovanni ha una credenza su quella persona particolare che è effettivamente il sindaco nel mondo attuale, credendo di quella persona che è corrotta, indipendentemente da come Giovanni la concettualizza o descrive.

Formalmente in semantica intensionale:

De dicto: M, w ⊨ B□(φ(ιx ψ(x))) sse per ogni w' tale che wR□w', se d è l'unico oggetto in D(w') che soddisfa ψ in w', allora M, w' ⊨ φ(d)

De re: M, w ⊨ B□(φ(d)) dove d = $[\![\iota x\ \psi(x)]\!]^w$ sse per ogni w' tale che wR□w', M, w' ⊨ φ(d)

L'oggetto d è fissato dal mondo attuale w nella lettura de re, mentre varia con i mondi doxastici w' nella lettura de dicto.

Questa distinzione ha conseguenze pratiche importanti per l'interpretazione di enunciati in contesti legali, giornalistici, e scientifici. Consideriamo l'enunciato da un articolo di giornale:

(12) "La polizia crede che il rapinatore della banca indossava una maschera"

Se interpretato de dicto, questo significa che la polizia crede la proposizione generale che chiunque abbia rapinato la banca indossava una maschera. Se interpretato de re, significa che c'è una persona particolare (identificata dalla polizia come il rapinatore) e la polizia crede di quella persona che indossava una maschera. La differenza è cruciale: la lettura de re presuppone che la polizia ha identificato un sospetto specifico, mentre la lettura de dicto lascia aperto chi sia il rapinatore.

La Semantica Formale dei Condizionali Controfattuali e la Logica delle Similarità tra Mondi

Il Problema Filosofico dei Condizionali Controfattuali

I condizionali controfattuali, anche chiamati condizionali soggettivi o contraffattuali, sono enunciati della forma "Se fosse stato il caso che A, allora sarebbe stato il caso che C", dove l'antecedente A è tipicamente falso o la sua verità è indeterminata, e l'enunciato esprime una dipendenza controfattuale tra A e C. Questi condizionali sono ubiquitari nel ragionamento ordinario, nella spiegazione causale, nell'attribuzione di responsabilità morale, e nell'analisi scientifica, eppure la loro semantica formale ha presentato sfide profonde che hanno occupato filosofi e logici per decenni.

Consideriamo esempi paradigmatici:

(1) "Se avessi studiato di più, avrei passato l'esame" (2) "Se Hitler fosse stato assassinato da bambino, la Seconda Guerra Mondiale non sarebbe accaduta" (3) "Se questo bicchiere fosse stato lasciato cadere, si sarebbe rotto" (4) "Se i canguri non avessero code, cadrebbero continuamente"

In ciascuno di questi casi, l'antecedente descrive una situazione controfattuale, diversa dalla realtà attuale, e il conseguente afferma cosa sarebbe stato il caso in quella circostanza alternativa. La questione teorica centrale è: quali sono le condizioni di verità di tali enunciati? Come determiniamo se un condizionale controfattuale è vero o falso?

La semantica del condizionale materiale della logica classica è notoriamente inadeguata per condizionali controfattuali. Il condizionale materiale $A \rightarrow C$ è definito come equivalente a $\neg A \vee C$, ovvero è vero automaticamente ogni volta che l'antecedente è falso. Questo genera i cosiddetti paradossi dell'implicazione materiale:

- Da un antecedente falso segue qualsiasi conseguente (ex falso quodlibet per il condizionale materiale)
- Qualsiasi condizionale con conseguente vero è automaticamente vero

Formalmente: $(\neg A) \vDash (A \rightarrow C)$ per qualsiasi C $C \vDash (A \rightarrow C)$ per qualsiasi A

Applicando questa semantica ai nostri esempi, otteniamo risultati assurdi. Dato che di fatto non ho studiato sufficientemente e l'antecedente di (1) è falso, il condizionale materiale implicherebbe che (1) è vero indipendentemente da cosa sarebbe successo se avessi

studiato. Ma anche il condizionale "Se avessi studiato di più, avrei fallito l'esame" sarebbe vero, e così ogni altro condizionale con quell'antecedente falso. Questo chiaramente non cattura il significato intuitivo dei condizionali controfattuali, che fanno affermazioni sostantive circa dipendenze tra situazioni ipotetiche.

Una seconda opzione, la logica modale stretta, interpreta il condizionale come necessità dell'implicazione materiale:

$A \mathrel{\Box\!\!\rightarrow} C =_{df} \Box(A \rightarrow C) =_{df} \Box(\neg A \vee C)$

Questo è vero sse in tutti i mondi possibili, o A è falso o C è vero. Ma anche questo è troppo forte: richiederebbe che il condizionale sia vero solo se C è una conseguenza logica o metafisicamente necessaria di A. Consideriamo (3): "Se questo bicchiere fosse lasciato cadere, si sarebbe rotto". È certamente possibile (anche se improbabile) che un bicchiere lasciato cadere non si rompa, magari perché cade su una superficie estremamente morbida o viene miracolosamente catturato. Quindi $\neg\Box$(Lasciato_cadere → Rotto), e il condizionale di necessità stretta sarebbe falso. Ma intuitivamente (3) sembra vero nelle circostanze normali.

La Semantica dei Mondi Possibili Più Vicini: L'Analisi di Stalnaker-Lewis

La soluzione teorica più influente al problema dei condizionali controfattuali fu sviluppata indipendentemente da Robert Stalnaker e David Lewis negli anni sessanta e settanta, basandosi sulla nozione di mondi possibili più vicini o più simili al mondo attuale. L'idea centrale è che un condizionale controfattuale "Se A allora C" è vero nel mondo attuale w_0 sse C è vero nei mondi possibili dove A è vero che sono più simili a w_0, ovvero nei mondi A-possibili che divergono minimalmente da w_0.

Formalmente, la semantica richiede una relazione di similarità comparativa tra mondi. Per ogni mondo w, definiamo una relazione $\leq_w$ (leggi: "almeno tanto simile a w quanto") che ordina parzialmente i mondi possibili rispetto alla loro similarità a w. Intuitivamente, $w_1 \leq_w w_2$ significa che w_1 è almeno tanto simile a w quanto lo è w_2, ovvero w_1 non è più dissimile da w di quanto lo sia w_2.

Stalnaker impone che per ogni mondo w e ogni proposizione A non-vuota, esiste un unico mondo A-più-vicino-a-w, denotato f(A, w), che è il mondo più simile a w tra tutti i mondi dove A è vero:

$f(A, w)$ = il mondo w' tale che: (i) $w' \vDash A$ (ii) $\forall w''\ [w'' \vDash A \rightarrow w' \leq_w w'']$

La condizione di verità per il condizionale controfattuale $A \mathrel{\Box\!\!\rightarrow} C$ diventa semplicemente:

$M, w \vDash (A \mathrel{\Box\!\!\rightarrow} C)$ sse $M, f(A,w) \vDash C$

Ovvero, "Se A allora C" è vero in w sse C è vero nel mondo A-più-vicino a w.

Lewis adotta un framework più generale che permette che possano non esistere mondi A-più-vicini singoli, ma piuttosto una sfera di mondi A-possibili ugualmente vicini che sono tutti più vicini di qualsiasi mondo A-possibile più lontano. La semantica lewisiana definisce:

M, w ⊨ (A □→ C) sse: (i) Non esistono mondi A-possibili accessibili da w, oppure (ii) Esiste una sfera S di mondi attorno a w tale che: - S contiene almeno un mondo A-possibile - Tutti i mondi A-possibili in S sono C-mondi - Nessun mondo fuori S è più vicino a w dei mondi in S

Intuitivamente: "Se A allora C" è vero sse o non ci sono mondi A-possibili (A è impossibile, e quindi il condizionale è vacuamente vero), oppure C è vero in tutti i mondi A-possibili nella regione di similarità più stretta attorno a w che contiene almeno un mondo A-possibile.

Formalizziamo più precisamente usando l'ordine di similarità. Diciamo che un insieme di mondi X è un A-permitting sphere attorno a w sse:

(i) ∃w' ∈ X tale che M, w' ⊨ A (ii) ∀w' ∈ X ∀w'' ∉ X [w' <w w''] (tutti i mondi in X sono più vicini a w di qualsiasi mondo fuori X)

La condizione di verità lewisiana diventa:

M, w ⊨ (A □→ C) sse: Non esistono mondi A-possibili, oppure ∃S [S è una A-permitting sphere attorno a w ∧ ∀w' ∈ S (M, w' ⊨ A → M, w' ⊨ C)]

Questa semantica valida una logica condizionale che evita i paradossi del condizionale materiale e del condizionale di necessità stretta. In particolare:

Modus Ponens Condizionale: A, (A □→ C) ⊨ C (Se A è vero e "se A allora C" è vero, allora C è vero)

Questo è valido perché se A è vero nel mondo attuale w, allora w stesso è il mondo A-più-vicino a w (dato che un mondo è sempre massimalmente simile a se stesso), quindi se il condizionale è vero, C deve essere vero in w.

Contrapposizione Controfattuale: (A □→ C) ⊨ (¬C □→ ¬A) Questa è INVALIDA nella logica di Lewis.

Per vedere perché, consideriamo un esempio. Supponiamo che il condizionale:

(5) "Se avessi colpito questa cartina con un fiammifero acceso, si sarebbe incendiata"

sia vero. La contrapposizione affermerebbe:

(6) "Se questa cartina non si fosse incendiata, non l'avrei colpita con un fiammifero acceso"

Ma (6) sembra falso: anche se la cartina non si è incendiata (fatto attuale), se l'avessi colpita con un fiammifero, si sarebbe incendiata. La non-incendiazione attuale non implica controfattualmente che non l'avrei colpita.

Il motivo logico del fallimento della contrapposizione è che il mondo ¬C-più-vicino potrebbe non essere identico al mondo A-più-vicino. I mondi dove la cartina non si incendia includono sia mondi dove non la colpisco con il fiammifero (molto simili al mondo attuale) sia mondi dove la colpisco ma la cartina è bagnata o il fiammifero non funziona (meno simili). Il mondo ¬C-più-vicino è plausibilmente uno del primo tipo, mentre il mondo A-più-vicino (dove la

colpisco) è del secondo tipo se consideriamo che la fiammata è il cambiamento minimale consistente con colpirla.

Transitività Controfattuale: (A □→ B), (B □→ C) ⊨ (A □→ C) Anche questa è generalmente INVALIDA.

Controesempio: Supponiamo che: (7) "Se Bizet e Verdi fossero stati connazionali, Bizet sarebbe stato italiano" (8) "Se Bizet fosse stato italiano, Bizet sarebbe stato italiano"

(7) è plausibilmente vero: dato che Verdi era italiano e Bizet francese, se dovessero essere connazionali, il cambiamento minimale sarebbe rendere Bizet italiano piuttosto che Verdi francese (dato che l'opera italiana è più centrale alla storia musicale di Verdi di quanto l'opera francese lo sia per Bizet, o per qualche altra considerazione pragmatica di naturalezza del cambiamento).

(8) è una tautologia condizionale, chiaramente vera.

Ma la conclusione della transitività sarebbe: (9) "Se Bizet e Verdi fossero stati connazionali, Bizet sarebbe stato italiano"

che è logicamente identica a (7), quindi vera. In questo caso la transitività sembra funzionare. Ma consideriamo un caso diverso:

(10) "Se Bizet e Verdi fossero stati connazionali, allora se Bizet non fosse stato italiano, Verdi sarebbe stato francese"

Questo sembra vero: dato che devono essere connazionali, se Bizet non è italiano, devono essere entrambi francesi.

Ma combinando (7) "Se connazionali, allora Bizet italiano" con (10), otterremmo per transitività:

(11) "Se Bizet e Verdi fossero stati connazionali, Verdi sarebbe stato francese"

che contraddice il nostro giudizio iniziale che il cambiamento minimale rende Bizet italiano, non Verdi francese. Il problema è che valutare condizionali nested (embedded) richiede considerare non i mondi più vicini al mondo attuale, ma i mondi più vicini ai mondi più vicini, e questa iterazione può produrre non-transitività.

Metriche di Similarità tra Mondi: Dimensioni e Ponderazioni

La semantica Stalnaker-Lewis lascia non specificato cosa esattamente costituisce "similarità" tra mondi. Quali dimensioni di confronto sono rilevanti? Come ponderiamo l'importanza relativa di diversi tipi di somiglianza?

Lewis propone alcuni principi generali per ordinare le dimensioni di similarità:

1. Evitamento di Miracoli Ampi: Mondi che condividono le stesse leggi naturali con il mondo attuale sono ceteris paribus più simili al mondo attuale di mondi con leggi diverse. I

"miracoli" (violazioni delle leggi naturali attuali) devono essere minimizzati in estensione spaziotemporale e significatività.

2. Massimizzazione di Matching di Fatti Particolari: Tra mondi che condividono le leggi, quelli che concordano su più fatti particolari sono più simili. Un mondo che diverge dal nostro soltanto riguardo a un singolo evento localizzato è più simile di uno che diverge riguardo a milioni di eventi.

3. Evitamento di Divergenze Gratuite: Una volta che un mondo diverge dal mondo attuale (per accomodare l'antecedente controfattuale), dovrebbe evolvere secondo le leggi naturali condivise senza divergenze addizionali non necessarie.

Lewis ordina queste dimensioni lessicograficamente: evitare miracoli ampi è più importante del matching di fatti particolari, che a sua volta è più importante dell'evitare divergenze piccole.

Formalmente, possiamo pensare alla relazione di similarità come determinata da una funzione di distanza multi-dimensionale:

$Dist(w, w') = \alpha \cdot D\_leggi(w, w') + \beta \cdot D\_fatti\_maggiori(w, w') + \gamma \cdot D\_fatti\_minori(w, w') + ...$

dove $\alpha >> \beta >> \gamma$ riflette l'ordinamento lessicografico delle priorità.

Applicando questi principi all'esempio canonico di Lewis:

(12) "Se Nixon avesse premuto il pulsante, sarebbe scoppiata una guerra nucleare"

Il mondo dove Nixon preme il pulsante più simile al mondo attuale è uno dove:

- Le leggi fisiche sono le stesse (i missili funzionano secondo le leggi dell'esplosione nucleare, eccetera)
- La storia del mondo fino al momento immediatamente prima che Nixon prema è identica al mondo attuale
- Al momento t_0, avviene un "piccolo miracolo" localizzato: Nixon preme il pulsante (divergendo dalla storia attuale dove non premette)
- Da t_0 in avanti, il mondo evolve secondo le leggi fisiche condivise: i sistemi di comando eseguono l'ordine, i missili vengono lanciati, la guerra nucleare inizia

Questo mondo richiede soltanto una divergenza miracolosa minima (la decisione di Nixon di premere), e poi tutto procede lawfully. Un mondo alternativo dove Nixon preme ma la guerra non scoppia richiederebbe miracoli molto più estesi e gravi: forse tutti i sistemi missilistici malfunzionano simultaneamente, o le esplosioni nucleari misteriosamente non accadono, violando leggi fisiche fondamentali. Questo mondo è molto più dissimile dal mondo attuale, quindi non è tra i mondi più vicini dove Nixon preme il pulsante.

Quindi (12) è vero: nei mondi più simili al nostro dove Nixon preme, la guerra scoppia.

Tuttavia, questa metrica di similarità è stata criticata per essere troppo oggettivista e fisica. Jonathan Bennett e altri hanno argomentato che la similarità rilevante per i condizionali ordinari include anche considerazioni pragmatiche, contestuali, e di salienza. In alcuni

contesti conversazionali, cambiare leggi naturali potrebbe essere considerato più "naturale" o "vicino" che cambiare molti fatti particolari, se il punto del condizionale è esplorare scenari metafisicamente differenti.

Per esempio:

(13) "Se la costante gravitazionale fosse stata più grande, l'universo sarebbe collassato presto dopo il Big Bang"

Questo condizionale è valutato cambiando una legge fondamentale della fisica (la costante gravitazionale), non attraverso un "miracolo" localizzato. La metrica di similarità per questo condizionale prioritizza la coerenza nomologica interna del mondo controfattuale (tutte le conseguenze fisiche seguono dalla legge alterata) sopra il matching con le leggi attuali.

Logica Controfattuale e Sistemi Assiomatici

La logica controfattuale può essere assiomatizzata formalmente. Il sistema VC (Variable Centered) di Lewis fornisce una assiomatizzazione completa rispetto alla semantica delle sfere che abbiamo descritto. Gli assiomi principali includono:

VC1: Tutti i teoremi della logica proposizionale classica (per enunciati senza operatori condizionali)

VC2: $(A \mathrel{\Box\!\!\rightarrow} A)$ Ogni condizionale con antecedente e conseguente identici è vero (tautologia condizionale)

VC3: $(A \mathrel{\Box\!\!\rightarrow} B) \rightarrow [(A \wedge C) \mathrel{\Box\!\!\rightarrow} B]$ Se "se A allora B" allora "se A e C allora B" (rafforzamento dell'antecedente)

VC4: $[(A \wedge B) \mathrel{\Box\!\!\rightarrow} C] \rightarrow [A \mathrel{\Box\!\!\rightarrow} (B \rightarrow C)]$ Principio di esportazione-importazione condizionale

VC5: $[(A \mathrel{\Box\!\!\rightarrow} B) \wedge (A \mathrel{\Box\!\!\rightarrow} C)] \rightarrow [A \mathrel{\Box\!\!\rightarrow} (B \wedge C)]$ Congiungibilità dei conseguenti con antecedente comune

VC6: $[(A \mathrel{\Box\!\!\rightarrow} C) \wedge (B \mathrel{\Box\!\!\rightarrow} C)] \rightarrow [(A \vee B) \mathrel{\Box\!\!\rightarrow} C]$ Principio di ragionamento per casi

Regole di inferenza:

Modus Ponens Classico: Da φ e $(\varphi \rightarrow \psi)$ inferire ψ

Equivalenza Sostitutiva: Se $\vdash (\varphi \leftrightarrow \psi)$, allora possiamo sostituire φ con ψ in qualsiasi contesto preservando teoremicità

Notevolmente, questo sistema NON include:

Transitività: $(A \mathrel{\Box\!\!\rightarrow} B) \wedge (B \mathrel{\Box\!\!\rightarrow} C) \rightarrow (A \mathrel{\Box\!\!\rightarrow} C)$ [NON valida]

Contrapposizione: $(A \mathrel{\Box\!\!\rightarrow} C) \rightarrow (\neg C \mathrel{\Box\!\!\rightarrow} \neg A)$ [NON valida]

Strengthening Inverso: [A □→ B] → [(A ∧ C) □→ B] è valida, ma l'inverso [(A ∧ C) □→ B] → [A □→ B] NON è valida

Il fatto che rafforzare l'antecedente è valido mentre indebolirlo non lo è riflette asimmetria fondamentale: se C è vero nei mondi A-più-vicini, allora C è anche vero nei mondi (A ∧ C)-più-vicini (che sono un sottoinsieme). Ma se C è vero nei mondi (A ∧ C)-più-vicini, potrebbe non essere vero nei mondi A-più-vicini che non soddisfano C.

Esempio: "Se avessi colpito questo fiammifero e fosse stato asciutto, si sarebbe acceso" (vero) Non implica: "Se avessi colpito questo fiammifero, si sarebbe acceso" (falso se il fiammifero è effettivamente bagnato)

Condizionali Nested e Iterazioni di Controfattualità

Una complicazione teorica profonda emerge quando consideriamo condizionali controfattuali nested, ovvero condizionali nel cui conseguente appare a sua volta un condizionale controfattuale:

(14) "Se fossimo stati in Australia, allora se avessimo visto un mammifero, probabilmente sarebbe stato un marsupiale"

Questo condizionale nested richiede valutazione in più passi: primo, consideriamo i mondi Australia-più-vicini (mondi dove siamo in Australia). Poi, all'interno di ciascuno di questi mondi, consideriamo i mondi vediamo-mammifero-più-vicini relativi a quel mondo. Il condizionale nested è vero sse nella maggior parte di questi mondi doppiamente controfattuali, il mammifero visto è un marsupiale.

Formalmente, la semantica è:

M, w ⊨ (A □→ (B □→ C)) sse per ogni w' ∈ Più-Vicini(w, A), M, w' ⊨ (B □→ C)

sse per ogni w' ∈ Più-Vicini(w, A), per ogni w'' ∈ Più-Vicini(w', B), M, w'' ⊨ C

Questa semantica iterata può produrre effetti controintuitivi. I mondi Più-Vicini(w', B) potrebbero non essere tra i mondi Più-Vicini(w, A ∧ B). Ovvero, il mondo (A ∧ B)-più-vicino-a-w potrebbe non essere il mondo B-più-vicino-al-mondo-A-più-vicino-a-w.

Questo fenomeno è chiamato **non-centering** del condizionale nested. Per illustrarlo, consideriamo l'esempio di Lewis:

(15) Mondo attuale: non piove, ma l'erba è bagnata (qualcuno l'ha innaffiata)

(16) "Se fosse piovuto, allora se l'erba fosse stata asciutta, non sarebbe piovuto molto"

Valutiamo questo nested condizionale:

- Primo passo: consideriamo mondi piove-più-vicini. In questi mondi, l'erba è tipicamente bagnata dalla pioggia (perché questo è l'effetto naturale della pioggia).

- Secondo passo: all'interno di questi mondi-pioggia, consideriamo mondi erba-asciutta-più-vicini. Questi sarebbero mondi dove piove ma la pioggia è leggera o breve, insufficiente a bagnare l'erba.
- Il condizionale nested afferma che in questi mondi doppiamente controfattuali (piove & erba-asciutta), non piove molto.

Questo sembra vero. Ma confrontiamo con:

(17) "Se fosse piovuto e l'erba fosse stata asciutta, non sarebbe piovuto molto"

Questo condizionale congiunto (non-nested) sembra equivalente intuitivamente a (16), ma la semantica formale può assegnare condizioni di verità diverse se i mondi (piove $\wedge$ erba-asciutta)-più-vicini al mondo attuale differiscono dai mondi erba-asciutta-più-vicini-ai-mondi-piove-più-vicini.

Questo problema ha generato dibattiti estesi circa se la semantica standard dei condizionali nested è corretta, e se potrebbe essere necessaria una semantica alternativa "centering" che garantisce:

Più-Vicini(Più-Vicini(w, A), B) = Più-Vicini(w, A $\wedge$ B)

Tuttavia, imporre centering genera altre complicazioni logiche e richiede modifiche sostanziali alla teoria.

Condizionali e Causalità: Teorie Controfattuali della Causazione

I condizionali controfattuali sono intimamente connessi all'analisi filosofica della causalità. Una delle teorie più influenti della causazione è la teoria controfattuale sviluppata da David Lewis, secondo cui un evento C causa un evento E sse:

(i) C ed E entrambi accadono (ii) Se C non fosse accaduto, E non sarebbe accaduto

Formalmente: Causa(C, E) =df C $\wedge$ E $\wedge$ ($\neg$C $\square\!\!\rightarrow$ $\neg$E)

Questo cattura l'intuizione che le cause sono condizioni necessarie nel senso controfattuale: rimuovendo la causa (mantenendo tutto il resto quanto più possibile simile), l'effetto non si verificherebbe.

Esempio: "Il cortocircuito ha causato l'incendio"

Analisi: Il cortocircuito accadde, l'incendio accadde, e se il cortocircuito non fosse accaduto (in circostanze simili per altri aspetti), l'incendio non si sarebbe verificato.

Tuttavia, questa analisi affronta problemi noti:

Problema della Preemption: Supponiamo che due processi causali indipendenti potrebbero ciascuno produrre l'effetto, ma uno "preempt" l'altro accadendo prima.

Esempio: Suzy lancia un sasso che colpisce una bottiglia rompendola. Billy lancia anche un sasso verso la bottiglia, ma il sasso di Suzy arriva prima e rompe la bottiglia, quindi il sasso di Billy non la colpisce mai.

Intuitivamente, il lancio di Suzy causò la rottura, ma il condizionale controfattuale "Se Suzy non avesse lanciato, la bottiglia non si sarebbe rotta" è FALSO (perché il sasso di Billy l'avrebbe rotta). Quindi la teoria controfattuale semplice non cattura questo caso di causazione.

Problema dell'Overdetermination Simmetrica: Due processi causali indipendenti ma simultanei producono lo stesso effetto, ciascuno sufficiente da solo.

Esempio: Due plotoni di esecuzione sparano simultaneamente a un condannato, ciascun plotone sparando abbastanza proiettili per causare morte. Il condannato muore.

Il lancio del plotone A causò la morte? Il condizionale "Se il plotone A non avesse sparato, il condannato non sarebbe morto" è falso (perché il plotone B l'avrebbe ucciso). Similmente per il plotone B. Quindi la teoria controfattuale implica che nessuno dei due plotoni causò la morte, il che è controintuitivo.

Per risolvere questi problemi, Lewis ha sviluppato una teoria più sofisticata basata su **dipendenza causale transitiva** e **catene causali**. L'idea è che anche se E non dipende controfattualmente direttamente da C, potrebbe esserci una catena di eventi intermedi C, D, E tale che D dipende controfattualmente da C e E dipende controfattualmente da D. In questo caso, C causa E attraverso la catena.

Formalmente: C causa E sse esiste una catena $C = C_0, C_1, C_2, ..., C_n = E$ tale che: per ogni $i < n$, C_{i+1} dipende controfattualmente da C_i

dove "D dipende controfattualmente da C" significa: $C \wedge D \wedge (\neg C \boxright \neg D) \wedge (C \boxright D)$

Questa teoria più elaborata può maneggiare molti casi problematici, ma continua ad affrontare contro-esempi sottili e rimane oggetto di raffinamento teorico continuo.

La Teoria degli Atti Linguistici e la Dimensione Performativa del Linguaggio

Il Programma Austiniano: Oltre il Fallacia Descrittivista

John Langshaw Austin, nelle sue celebri conferenze William James del millenovecentocinquantacinque pubblicate postume come "How to Do Things with Words" nel millenovecentosessantadue, rivoluzionò la filosofia del linguaggio sfidando quella che chiamò la "fallacia descrittivista" (descriptive fallacy), ovvero l'assunzione dominante nella tradizione filosofica analitica che la funzione primaria o esclusiva del linguaggio sia descrivere stati di cose e che quindi il compito centrale della semantica sia specificare condizioni di verità per enunciati dichiarativi.

Austin osservò che molti enunciati nel linguaggio ordinario non descrivono affatto il mondo e non sono valutabili come veri o falsi, ma piuttosto **eseguono azioni** (perform actions) attraverso la loro enunciazione. Consideriamo i seguenti esempi:

(1) "Vi dichiaro marito e moglie" (pronunciato da un celebrante durante una cerimonia matrimoniale) (2) "Prometto di restituirti il libro domani" (3) "Battezzo questa nave HMS Queen Elizabeth" (pronunciato mentre si rompe una bottiglia contro lo scafo) (4) "Mi scuso per il mio ritardo" (5) "Ti ordino di lasciare la stanza"

Nessuno di questi enunciati descrive una situazione preesistente che potrebbe essere vera o falsa. Invece, pronunciare questi enunciati nelle circostanze appropriate **costituisce** l'esecuzione delle azioni di sposare, promettere, battezzare, scusarsi, ordinare. Austin chiamò tali enunciati **performativi** (performatives) in contrasto con gli enunciati **constativi** (constatives) che si limitano a descrivere o riportare fatti.

La caratteristica distintiva dei performativi è che la loro enunciazione felice (felicitous utterance) non consiste nell'essere vera ma nel soddisfare certe **condizioni di felicità** (felicity conditions) che determinano se l'atto linguistico è riuscito o fallito, appropriato o inappropriato. Austin articola quattro condizioni generali di felicità per performativi:

Condizione A.1 (Esistenza di Procedura Convenzionale): Deve esistere una procedura convenzionale accettata che ha un certo effetto convenzionale, e la procedura deve includere l'enunciamento di certe parole da parte di certe persone in certe circostanze.

Condizione A.2 (Appropriatezza dei Partecipanti e Circostanze): Le persone e le circostanze particolari del caso devono essere appropriate per l'invocazione della procedura particolare invocata.

Condizione B.1 (Esecuzione Corretta): La procedura deve essere eseguita da tutti i partecipanti correttamente.

Condizione B.2 (Esecuzione Completa): La procedura deve essere eseguita completamente.

Condizione Γ.1 (Sincerità dei Pensieri): Dove, come spesso, la procedura è disegnata per l'uso da parte di persone che hanno certi pensieri o sentimenti, o per l'inaugurazione di certa condotta conseguente da parte di qualche partecipante, allora una persona che partecipa e quindi invoca la procedura deve effettivamente avere quei pensieri o sentimenti.

Condizione Γ.2 (Condotta Conseguente): I partecipanti devono effettivamente condursi conseguentemente in seguito.

Violazioni delle condizioni A e B rendono l'atto nullo o vuoto (**misfires**): l'atto semplicemente non si realizza. Violazioni delle condizioni Γ rendono l'atto **abusivo** (**abuses**): l'atto si realizza ma in modo difettoso o insincero.

Esempi di misfires:

(6) "Vi dichiaro marito e moglie" pronunciato da qualcuno che non è un celebrante autorizzato → Violazione di A.2: il parlante non ha l'autorità appropriata

(7) "Battezzo questa nave..." pronunciato senza rompere la bottiglia → Violazione di B.1: la procedura non è eseguita completamente con le azioni richieste

(8) "Ti nomino Cavaliere" pronunciato da un cittadino ordinario (non dalla Regina o autorità delegata) → Violazione di A.2: mancanza di autorità istituzionale

Esempi di abuses:

(9) "Prometto di restituirti il libro" pronunciato senza intenzione di restituire il libro → Violazione di Γ.1: mancanza di sincerità

(10) "Mi congratulo per il tuo successo" pronunciato senza provare alcuna congratulazione genuina → Violazione di Γ.1: insincerità

Tuttavia, Austin riconosce progressivamente che la distinzione iniziale tra performativi e constativi è instabile. Anche gli enunciati apparentemente constativi, puramente descrittivi, possono essere analizzati come eseguenti atti linguistici. Consideriamo:

(11) "Il gatto è sulla stuoia"

Questo sembra un paradigma di enunciato constativo, semplicemente descrivente uno stato di cose. Ma pronunciare questo enunciato in un contesto conversazionale normale costituisce anche l'esecuzione dell'atto di **asserire** o **affermare** che il gatto è sulla stuoia, e questo atto di asserzione ha le sue proprie condizioni di felicità: il parlante dovrebbe essere in posizione epistemica per fare l'asserzione (credere che il gatto sia sulla stuoia, avere evidenza adeguata), l'asserzione dovrebbe essere rilevante per il discorso in corso, eccetera.

Questa osservazione porta Austin a sostituire la distinzione performativo/constativo con una teoria più generale degli **atti linguistici** (speech acts) che distingue tre livelli o dimensioni di azione eseguiti simultaneamente in ogni enunciamento linguistico completo:

Atto Locutorio (Locutionary Act): L'atto di pronunciare parole con un certo significato sintattico e semantico. Questo include:

- Atto fonetico: produrre certi suoni
- Atto fatico: produrre certi vocaboli o parole in conformità a certe regole grammaticali
- Atto retico: usare quei vocaboli con un certo senso e riferimento approssimativamente determinati

Atto Illocutorio (Illocutionary Act): L'atto **eseguito nel** (performed in) pronunciare l'enunciato, ovvero l'azione sociale o istituzionale che viene effettuata attraverso l'enunciamento: asserire, promettere, ordinare, chiedere, avvertire, eccetera.

Atto Perlocutorio (Perlocutionary Act): L'atto eseguito **mediante** (performed by) l'enunciamento, ovvero gli effetti o conseguenze causali che l'enunciamento produce sull'ascoltatore o sulla situazione: convincere, persuadere, allarmare, annoiare, divertire, eccetera.

Formalizziamo questa tripartizione attraverso un esempio concreto:

(12) Un poliziotto dice a qualcuno "C'è un toro là dietro"

- Atto Locutorio: Il poliziotto pronuncia le parole "C'è un toro là dietro" riferendosi a una particolare direzione e a un toro particolare, con il significato letterale "esiste un toro in quella locazione"

- Atto Illocutorio: Il poliziotto **avverte** l'ascoltatore del pericolo (questo è l'atto illocutorio: un warning/avvertimento)

- Atto Perlocutorio: L'ascoltatore si allontana dalla direzione indicata, evitando così il pericolo (l'atto perlocutorio: causare che l'ascoltatore eviti il pericolo)

Crucialmente, lo stesso atto locutorio può costituire atti illocutori diversi in contesti diversi:

(13) "C'è un toro là dietro" pronunciato da una guida turistica durante un safari fotografico

- Atto Illocutorio: **Informare** i turisti della presenza del toro per scopi di osservazione (non avvertimento di pericolo)

(14) "C'è un toro là dietro" pronunciato in una discussione sull'allevamento di bovini

- Atto Illocutorio: **Asserire** un fatto circa la distribuzione spaziale degli animali

L'atto illocutorio è determinato non puramente dal contenuto semantico dell'enunciato ma dall'**intenzione comunicativa** del parlante (ciò che il parlante intende fare attraverso l'enunciamento) in relazione alle convenzioni linguistiche e sociali che governano quell'uso del linguaggio.

La Tassonomia Searleana degli Atti Illocutori

John Searle, studente e collaboratore di Austin, sviluppò sistematicamente la teoria degli atti linguistici in direzioni più formalmente rigorose. Nel suo articolo del millenovecentosessantanove "A Taxonomy of Illocutionary Acts", Searle propone una classificazione sistematica degli atti illocutori basata su diverse dimensioni teoriche:

Dimensione 1: Scopo Illocutorio (Illocutionary Point) Lo scopo o obiettivo che definisce il tipo di atto. Per esempio:

- Assertivi: impegnare il parlante alla verità della proposizione espressa
- Direttivi: tentare di far fare qualcosa all'ascoltatore
- Commissivi: impegnare il parlante a un'azione futura
- Espressivi: esprimere stati psicologici del parlante
- Dichiarativi: creare corrispondenza tra contenuto proposizionale e realtà attraverso la dichiarazione stessa

Dimensione 2: Direzione di Adattamento (Direction of Fit) La relazione tra parole e mondo che l'atto cerca di stabilire:

- Parole-al-Mondo (word-to-world): le parole dovrebbero corrispondere al mondo (caratteristica di assertivi: descrivo il mondo come è) Simbolo: ↓

- Mondo-alle-Parole (world-to-word): il mondo dovrebbe cambiare per corrispondere alle parole (caratteristica di direttivi e commissivi: faccio richieste che cambieranno il mondo) Simbolo: ↑

- Bidirezionale: simultaneamente parole-al-mondo e mondo-alle-parole (caratteristica di dichiarativi: dichiarando qualcosa, lo rendo vero) Simbolo: ↕

- Nessuna direzione (null): né le parole devono corrispondere al mondo né il mondo deve cambiare (caratteristica di espressivi: semplicemente esprimo stati) Simbolo: ∅

Dimensione 3: Condizioni di Sincerità (Sincerity Conditions) Lo stato psicologico espresso dall'atto:

- Assertivi esprimono credenza (belief) che p
- Direttivi esprimono desiderio/volontà (desire/want) che l'ascoltatore faccia A
- Commissivi esprimono intenzione (intention) di fare A
- Espressivi esprimono vari stati emotivi o attitudinali
- Dichiarativi: nessuna condizione di sincerità speciale richiesta

Dimensione 4: Forza Illocutoria (Illocutionary Force) L'intensità o grado con cui lo scopo illocutorio è perseguito. Per esempio, nella classe dei direttivi:

- Suggerire (bassa forza)
- Richiedere (forza media)
- Ordinare (alta forza)
- Implorare (alta forza con connotazione di supplica)

Basandosi su queste dimensioni, Searle identifica cinque classi fondamentali:

Assertivi (Representatives/Assertives)

Scopo Illocutorio: Impegnare il parlante (in vari gradi) alla verità della proposizione espressa

Direzione di Adattamento: Parole↓Mondo

Condizione di Sincerità: Credenza(parlante, p)

Esempi: asserire, affermare, concludere, dichiarare, descrivere, diagnosticare, insistere, predire, riportare, teorizzare

Condizioni di Felicità Specifiche:

- Condizione Preparatoria: Il parlante ha evidenza (ragioni, basi) per la verità di p

- Condizione Essenziale: L'enunciamento conta come un'assunzione di impegno alla verità di p
- Condizione di Sincerità: Il parlante crede p

Formalizzazione parziale per "asserire":

ASSERT(S, p, H) è felice sse:
(i) S crede p [sincerità]
(ii) S ha evidenza per p [preparatoria]
(iii) Non è ovvio a H che S crede p [preparatoria - pertinenza]
(iv) L'enunciamento di S conta come un'assunzione di responsabilità per la verità di p [essenziale]

11.2.2 Direttivi (Directives)

Scopo Illocutorio: Tentare di far fare qualcosa all'ascoltatore

Direzione di Adattamento: Mondo↑Parole

Condizione di Sincerità: Desiderio/Volontà(parlante, H fa A)

Esempi: chiedere, comandare, ordinare, pregare, supplicare, implorare, invitare, permettere, consigliare, raccomandare

Condizioni di Felicità per "ordinare":

ORDER(S, H, A) è felice sse:
(i) S vuole che H faccia A [sincerità]
(ii) S ha autorità su H rispetto a fare A [preparatoria]
(iii) Non è ovvio che H farà A nel corso normale degli eventi [preparatoria]
(iv) S può causare che H faccia A attraverso l'ordine [preparatoria]
(v) L'enunciamento conta come un tentativo di far fare A a H [essenziale]

Differenza tra diversi tipi di direttivi risiede principalmente nella relazione di autorità e nella forza:

- **Ordinare**: S ha autorità su H, alta forza
- **Richiedere**: relazione più egualitaria, forza media
- **Pregare/Implorare**: S in posizione inferiore, fa appello alla benevolenza di H
- **Consigliare**: orientato al beneficio di H, non all'interesse di S

Commissivi (Commissives)

Scopo Illocutorio: Impegnare il parlante a un'azione futura

Direzione di Adattamento: Mondo↑Parole (il parlante farà cambiare il mondo)

Condizione di Sincerità: Intenzione(parlante, fare A)

Esempi: promettere, giurare, garantire, contrattare, accettare, acconsentire, offrire

Condizioni di Felicità per "promettere":

PROMISE(S, H, A) è felice sse:
(i) S intende fare A [sincerità]
(ii) H preferisce che S faccia A piuttosto che non-A [preparatoria]
(iii) Non è ovvio a S e H che S farebbe A nel corso normale degli eventi [preparatoria]
(iv) S è capace di fare A [preparatoria]
(v) L'enunciamento conta come un'assunzione di obbligo da parte di S di fare A [essenziale]

Una caratteristica distintiva delle promesse è che creano **obbligazioni normative**: dopo aver promesso, S ha un obbligo morale/sociale di fare A, indipendentemente da se S continua a voler fare A.

Espressivi (Expressives)

Scopo Illocutorio: Esprimere lo stato psicologico del parlante circa uno stato di cose specificato nel contenuto proposizionale

Direzione di Adattamento: Nessuna (∅)

Condizione di Sincerità: Vari stati psicologici (gioia, dolore, piacere, dispiacere, eccetera)

Esempi: ringraziare, congratularsi, scusarsi, condolersi, dare il benvenuto, esprimere gratitudine

Condizioni per "scusarsi":

APOLOGIZE(S, H, A) è felice sse:
(i) S ha fatto A (o è responsabile per A)
(ii) A è contrario agli interessi di H
(iii) S crede che A è contrario agli interessi di H
(iv) S prova dispiacere/rammarico per A [sincerità]
(v) L'enunciamento conta come espressione di dispiacere per A

Gli espressivi non cercano di descrivere il mondo (non hanno direzione parole-al-mondo) né di cambiarlo (non hanno direzione mondo-alle-parole), ma semplicemente manifestano pubblicamente uno stato interno del parlante.

Dichiarativi (Declarations)

Scopo Illocutorio: Creare una corrispondenza tra contenuto proposizionale e realtà attraverso la dichiarazione stessa

Direzione di Adattamento: Bidirezionale (↕)

Condizione di Sincerità: Nessuna speciale (l'atto è tipicamente legato a istituzioni)

Esempi: dichiarare guerra, battezzare, licenziare, nominare, sposare, assolvere, scomunicare

Caratteristica Distintiva: I dichiarativi richiedono **istituzioni extralinguistiche** che conferiscono autorità al parlante di effettuare cambiamenti di status attraverso dichiarazioni.

DECLARE_MARRIED(S, H_1, H_2) è felice sse:
(i) S ha l'autorità istituzionale di sposare persone [preparatoria critica]
(ii) H_1 e H_2 soddisfano i requisiti legali per sposarsi [preparatoria]
(iii) La cerimonia è eseguita correttamente [preparatoria]
(iv) Mediante la dichiarazione, S crea lo stato di H_1 e H_2 essendo sposati [essenziale]

La caratteristica peculiare è che la dichiarazione felice **crea la realtà** che descrive: dicendo "Vi dichiaro marito e moglie" nelle circostanze appropriate, il celebrante rende vero che le persone sono sposate. Le parole si adattano al mondo (descrivono il nuovo status) e simultaneamente il mondo si adatta alle parole (il nuovo status viene creato).

Atti Linguistici Indiretti e Implicature Convenzionali

Un fenomeno pragmatico pervasivo è l'uso di **atti linguistici indiretti** (indirect speech acts), dove il parlante esegue un atto illocutorio attraverso l'esecuzione di un altro. Il caso paradigmatico riguarda le richieste cortesi:

(15) "Puoi passarmi il sale?"

Questo ha la forma sintattica di una domanda sì/no circa le capacità dell'ascoltatore (un atto illocutorio diretto di richiesta di informazione). Ma in contesti normali di pasto condiviso, è usato idiomaticamente per **richiedere** cortesemente che l'ascoltatore passi il sale (atto illocutorio indiretto).

Searle analizza come gli ascoltatori inferiscono l'atto indiretto dall'atto diretto:

Passo 1: L'ascoltatore riconosce che è stata eseguita una domanda circa capacità (atto diretto primario)

Passo 2: L'ascoltatore riconosce che chiedere circa capacità dell'ascoltatore di fare A è plausibilmente rilevante soltanto se il parlante vuole che A venga fatto

Passo 3: L'ascoltatore riconosce che il contesto (pasto condiviso) rende saliente il desiderio del parlante di avere il sale

Passo 4: L'ascoltatore riconosce che le condizioni preparatorie per richiedere che l'ascoltatore passi il sale sono soddisfatte (l'ascoltatore è capace, non lo farebbe altrimenti, eccetera)

Passo 5: L'ascoltatore inferisce che il parlante sta richiedendo indirettamente che passi il sale

Passo 6: L'ascoltatore riconosce che il parlante riconosce che l'ascoltatore può fare questa inferenza

Conclusione: L'enunciamento conta come una richiesta indiretta cortese di passare il sale

Questo ragionamento inferenziale può essere formalizzato utilizzando principi conversazionali griceani (che esamineremo più approfonditamente successivamente). L'idea cruciale è che gli atti indiretti sfruttano convenzioni pragmatiche: certe forme linguistiche sono convenzionalmente associate a certi atti indiretti in virtù di regolarità pragmatiche consolidate nella comunità linguistica.

Esistono gradi di convenzionalizzazione:

Atti Indiretti Non-Convenzionalizzati: L'inferenza richiede ragionamento sostanziale circa gli scopi del parlante nel contesto particolare.

Esempio: "È caldo qui" pronunciato in un ufficio chiuso → Possibile richiesta indiretta di aprire una finestra, ma richiede inferenza contestuale

Atti Indiretti Semi-Convenzionalizzati: Certe formule linguistiche sono usate idiomaticamente per certi atti indiretti ma mantengono un collegamento inferenziale trasparente.

Esempi:

- "Puoi X?" = richiesta cortese di fare X
- "Potresti X?" = richiesta più cortese
- "Mi passeresti X?" = richiesta
- "Ti dispiacerebbe X?" = richiesta molto cortese

Atti Indiretti Completamente Convenzionalizzati (Idiomi Performativi): L'associazione tra forma linguistica e atto indiretto è completamente convenzionalizzata al punto che non richiede inferenza.

Esempi:

- "Come va?" = saluto, non genuina domanda circa benessere
- "Prego" dopo "grazie" = acknowledgment cortese, non invito
- "Ci vediamo" come congedo = saluto, non letteralmente piano di incontrarsi

Formalmente, possiamo modellare atti indiretti come una funzione composizionale:

Forza_Illocutoria_Totale(U) = Forza_Diretta(U) $\oplus$ Forza_Indiretta(U, Contesto)

dove:
Forza_Diretta(U) = forza illocutoria determinata dalla forma sintattica e markers espliciti
Forza_Indiretta(U, C) = forza illocutoria inferita da convenzioni pragmatiche + ragionamento contestuale
$\oplus$ = operatore di combinazione (tipicamente, la forza indiretta domina quando presente)

Marcatori di Forza Illocutoria e Indicatori di Modo

Nel linguaggio naturale, la forza illocutoria è segnalata attraverso diversi mezzi linguistici:

1. Verbi Performativi Espliciti: Verbi che, quando usati alla prima persona singolare presente indicativo, eseguono l'atto che nominano:

- "Ti prometto che..." → esegue una promessa
- "Ti ordino di..." → esegue un ordine
- "Mi scuso per..." → esegue una scusa
- "Ti avverto che..." → esegue un avvertimento

Questi funzionano come **indicatori di forza illocutoria esplicita** (explicit illocutionary force indicating devices, IFIDs).

2. Modo Verbale: Molte lingue usano morfologia verbale per indicare forza illocutoria:

- Modo Indicativo: tipicamente usato per assertivi "Giovanni parte domani"
- Modo Imperativo: tipicamente usato per direttivi "Parti immediatamente!"
- Modo Congiuntivo: esprime possibilità, desiderio, dubbio "Che egli parta!" (desiderativo/esortativo) "Supponiamo che parta" (ipotetico)
- Modo Condizionale: esprime richieste cortesi, proposte "Partiresti con me?" (richiesta cortese)

3. Ordine delle Parole e Inversione: In molte lingue, l'inversione soggetto-verbo segnala interrogativi:

Italiano: "Giovanni parte" (assertivo) vs "Parte Giovanni?" (interrogativo)

4. Intonazione e Prosodia: I pattern intonazionali distinguono atti illocutori:

- Intonazione ascendente: interrogativi "Parti↗?"
- Intonazione discendente: assertivi "Parti↘."
- Intonazione enfatica: esclamativi o enfasi "CHE bellezza!"

5. Particelle Modali: Molte lingue usano particelle specializzate per modulare forza illocutoria:

Tedesco:

- "doch" (rafforza asserzioni)
- "ja" (cerca conferma)
- "mal" (attenua richieste)

Giapponese:

- "ne" (cerca conferma)
- "yo" (assertivo forte)
- "ka" (interrogativo)

6. Marcatori Illocutivi Lessicali: Avverbi e espressioni che modificano forza:

- "Francamente, ..." (segnala franchise)
- "Sinceramente, ..." (segnala sincerità)
- "Per favore, ..." (segnala richiesta cortese)
- "Mi permetto di suggerire..." (attenua suggerimento)

Formalmente, possiamo rappresentare la forza illocutoria come una struttura composizionale:

$F = \langle \text{Tipo}, \text{Forza}, \text{Modalità}, \text{Condizioni} \rangle$

dove:
$\text{Tipo} \in \{\text{Assertivo}, \text{Direttivo}, \text{Commissivo}, \text{Espressivo}, \text{Dichiarativo}\}$
$\text{Forza} \in \mathbb{R}^+$ (intensità numerica)
$\text{Modalità} = \{\text{marcatori linguistici}\}$
$\text{Condizioni} = \{\text{condizioni di felicità}\}$

Un enunciato completo U ha una struttura:

$U = \langle C\_prop, F, \text{Contesto} \rangle$

dove:
C_prop = contenuto proposizionale (ciò su cui verte l'atto)
F = forza illocutoria
Contesto = parametri contestuali rilevanti

L'interpretazione pragmatica completa richiede determinare tutti e tre i componenti dall'input linguistico e dal contesto comunicativo.

La Teoria della Pertinenza di Sperber e Wilson: Un Approccio Cognitivo alla Comunicazione

Fondamenti Cognitivi della Comunicazione Umana: Ostensione e Inferenza

Dan Sperber e Deirdre Wilson, nella loro opera fondamentale "Relevance: Communication and Cognition" del millenovecentoottantasei, propongono una ricostruzione radicale dei fondamenti teorici della pragmatica linguistica, spostando l'enfasi da principi conversazionali cooperativi (come nella teoria griceana) a principi cognitivi di economia inferenziale basati

sulla pertinenza. La teoria della pertinenza (Relevance Theory) sostiene che la comunicazione umana è fondamentalmente un processo ostensivo-inferenziale: i parlanti producono stimoli ostensivi (enunciati e comportamenti che manifestano apertamente un'intenzione comunicativa) e gli ascoltatori inferiscono le intenzioni comunicative del parlante attraverso processi di ragionamento pragmatico guidati da aspettative di pertinenza ottimale.

Il punto di partenza è una distinzione cruciale tra due modelli fondamentalmente diversi di comunicazione:

Modello del Codice (Code Model): La comunicazione consiste nella codifica di messaggi in segnali da parte del parlante e nella decodifica di quei segnali in messaggi da parte dell'ascoltatore. Questo modello tratta la comunicazione come trasmissione di informazione attraverso un codice condiviso, analogamente alla comunicazione telefonica o telegrafica.

Formalmente:

Parlante: Messaggio M → Codifica(M) → Segnale S
Canale: trasmissione di S
Ascoltatore: Segnale S → Decodifica(S) → Messaggio M

Il modello del codice funziona perfettamente per sistemi di comunicazione artificiali e altamente regolamentati (codice Morse, linguaggi di programmazione, protocolli di comunicazione digitale), ma Sperber e Wilson argomentano che è fondamentalmente inadeguato per la comunicazione linguistica umana ordinaria, per diverse ragioni:

Sottodeterminazione Semantica: Come abbiamo esaminato estensivamente nei capitoli precedenti sul arricchimento pragmatico, il contenuto semantico codificato linguisticamente sottodetermina radicalmente il messaggio comunicato. L'enunciato "Ho fatto colazione" codifica soltanto che il parlante ha eseguito un'azione di fare-colazione in qualche momento passato, ma comunica tipicamente che il parlante ha fatto colazione nella mattinata dello stesso giorno. Questa specificazione temporale non è codificata ma inferita pragmaticamente.

Ambiguità Non-Risolte dalla Codifica: Espressioni linguistiche sono sistematicamente ambigue a livelli lessicali, sintattici e referenziali, e la risoluzione di queste ambiguità richiede inferenza pragmatica che va oltre la decodifica.

Esempio: "Il pollo è pronto da mangiare"

- Lettura 1: Il pollo (cibo) è cucinato e può essere mangiato
- Lettura 2: Il pollo (animale) è affamato e pronto a mangiare

La selezione tra queste letture non è determinata dal codice linguistico ma richiede ragionamento inferenziale basato su conoscenza del mondo e contesto.

Implicature e Contenuto Implicito: Una porzione sostanziale di ciò che viene comunicato attraverso enunciati non è codificato esplicitamente ma è implicato pragmaticamente.

Esempio: A: "Vuoi venire al cinema stasera?" B: "Devo studiare per l'esame domani"

B non ha codificato "No, non verrò al cinema", ma questa risposta è comunicata implicitamente attraverso implicatura conversazionale.

Modello Ostensivo-Inferenziale (Ostensive-Inferential Model): La comunicazione consiste nel rendere manifesta un'intenzione di rendere manifesta informazione. Il comunicatore produce un stimolo ostensivo che rende mutuamente manifesto al comunicatore e al destinatario che il comunicatore intende rendere manifesto un insieme di assunzioni. Il destinatario inferisce le intenzioni informative del comunicatore basandosi su questo stimolo ostensivo e su aspettative di pertinenza.

Formalmente:

Comunicatore: Intende rendere manifesto I (informazione)
→ Produce stimolo ostensivo S che manifesta questa intenzione

Destinatario: Percepisce S
→ Riconosce che S è ostensivo (manifesta intenzione comunicativa)
→ Inferisce I basandosi su S + Contesto + Aspettative di Pertinenza
→ Riconosce che il Comunicatore intendeva che I fosse inferito

La nozione di **manifestezza** (manifestness) è centrale nella teoria. Un'assunzione è manifesta a un individuo in un tempo dato sse l'individuo è capace in quel tempo di rappresentarsi mentalmente quell'assunzione e accettarla come vera o probabilmente vera. Il **ambiente cognitivo** di un individuo è l'insieme di tutte le assunzioni che sono manifeste a quell'individuo.

Definizione formale:

Manifesto(A, I, t) =df
I è capace al tempo t di rappresentarsi mentalmente l'assunzione A
∧
I è capace di inferire o percepire evidenza che rende A vera o probabilmente vera

AmbienteCognitivo(I, t) = {A | Manifesto(A, I, t)}

Un fatto è **mutuamente manifesto** a un insieme di individui $\{I_1, I_2, ..., I_□\}$ sse è manifesto a ciascuno di loro che quel fatto è manifesto a tutti loro.

MutuamenteManifeosto(A, $\{I_1, ..., I_□\}$, t) =df
$\forall i \in \{1..n\}$: Manifesto(A, I_i, t)
∧
$\forall i \in \{1..n\}$: Manifesto(MutuamenteManifeasto(A, $\{I_1, ..., I_□\}$, t), I_i, t)

Questa è una definizione ricorsiva che cattura l'idea di conoscenza comune (common knowledge): tutti sanno A, tutti sanno che tutti sanno A, tutti sanno che tutti sanno che tutti sanno A, e così via all'infinito.

La comunicazione ostensiva crea **mutua manifestezza** tra comunicatore e destinatario dell'intenzione comunicativa del comunicatore. Quando un parlante produce ostensivamente un enunciato, rende mutuamente manifesto che: (i) Il parlante ha un'intenzione informativa (vuole rendere manifeste certe assunzioni) (ii) Il parlante ha un'intenzione comunicativa (vuole rendere manifesta questa intenzione informativa) (iii) L'ascoltatore dovrebbe inferire le assunzioni che il parlante intende comunicare

Il Principio Comunicativo di Pertinenza e l'Euristica di Comprensione

Il cuore teorico della Relevance Theory è il **Principio Comunicativo di Pertinenza** (Communicative Principle of Relevance), che afferma che ogni atto di comunicazione ostensiva comunica automaticamente una presunzione della propria pertinenza ottimale. Questa presunzione guida il processo interpretativo dell'ascoltatore, permettendo inferenze efficienti delle intenzioni del parlante senza bisogno di stipulazioni esplicite di massime conversazionali come nella teoria griceana.

Prima di formulare il principio precisamente, dobbiamo definire rigorosamente la nozione di **pertinenza** (relevance). Sperber e Wilson propongono che la pertinenza è una proprietà di input informativi (stimoli esterni o rappresentazioni mentali interne) relativamente a un individuo e a un tempo, e che ha due dimensioni:

Dimensione 1: Effetti Cognitivi Positivi (Positive Cognitive Effects) Un input I è pertinente per un individuo S sse processare I nel contesto cognitivo di S produce effetti cognitivi positivi. Gli effetti cognitivi positivi includono:

(a) **Implicazioni Contestuali**: Nuove conclusioni derivabili dalla combinazione di I con assunzioni già presenti nel contesto cognitivo ma non derivabili senza I

Formalmente:

ImplicazioniContestuali(I, C) =
$\{A \mid (C \cup \{I\}) \vdash A \wedge C \nvdash A \wedge \{I\} \nvdash A\}$

Ovvero, implicazioni contestuali sono assunzioni che seguono dall'unione di I con il contesto C ma non seguono da C da solo né da I da solo.

Esempio:

Contesto C: {Se piove, la strada sarà bagnata}
Input I: {Sta piovendo}
Implicazione contestuale: {La strada sarà bagnata}

(b) **Rafforzamento di Assunzioni Esistenti**: I fornisce evidenza addizionale per assunzioni già presenti in C, aumentandone la forza epistemica

Rafforzamento(A, I, C) =
  ΔForza(A) dove A ∈ C ∧ I fornisce evidenza per A

Esempio:

Contesto C: {Giovanni è probabilmente a casa (forza 0,6)}
Input I: {La luce della casa di Giovanni è accesa}
Rafforzamento: {Giovanni è a casa (forza 0,85)}

(c) **Eliminazione o Indebolimento di Assunzioni**: I contraddice assunzioni in C, portando alla loro revisione o eliminazione

Contraddizione(I, C) =
  {A ∈ C | I ⊢ ¬A} (assunzioni in C contraddette da I)

Esempio:

Contesto C: {Tutti i cigni sono bianchi}
Input I: {Ho osservato un cigno nero in Australia}
Eliminazione: {Tutti i cigni sono bianchi} viene rimosso o indebolito

La quantità totale di effetti cognitivi positivi di I in C può essere formalmente definita come:

EffettiCognitivi(I, C) =
  |ImplicazioniContestuali(I, C)|
  + Σ[A∈C] Rafforzamento(A, I, C)
  + Σ[A∈C] Indebolimento(A, I, C)

Dimensione 2: Sforzo di Elaborazione (Processing Effort) L'sforzo cognitivo richiesto per processare I nel contesto C include:

- Sforzo di accesso percettivo o di memoria per I
- Sforzo di recupero di assunzioni contestuali rilevanti da C
- Sforzo inferenziale per derivare effetti cognitivi

Formalmente, possiamo modellare lo sforzo come una funzione delle risorse cognitive utilizzate:

Sforzo(I, C) =
  α · Complessità_Percettiva(I)
  + β · Complessità_Recupero_Contestuale(C)
  + γ · Complessità_Inferenziale(I, C)

dove α, β, γ sono parametri di ponderazione.

La **pertinenza** è definita come una funzione crescente di effetti cognitivi e decrescente di sforzo:

Pertinenza(I, C) = f(EffettiCognitivi(I, C), Sforzo(I, C))

Una formalizzazione possibile è il rapporto:

Pertinenza(I, C) = EffettiCognitivi(I, C) / Sforzo(I, C)

Altre cose essendo uguali:

- Maggiori effetti cognitivi → maggiore pertinenza
- Minore sforzo → maggiore pertinenza

Principio Comunicativo di Pertinenza (formulazione completa):

Ogni atto di comunicazione ostensiva comunica una presunzione della propria pertinenza ottimale, ovvero:

(a) Lo stimolo ostensivo è sufficientemente pertinente da meritare lo sforzo di elaborazione del destinatario

(b) Lo stimolo ostensivo è il più pertinente compatibile con le abilità e le preferenze del comunicatore

Formalmente:

∀ S_ostensivo prodotto dal Comunicatore C:

(a) ∃ I [Interpretazione(S, I) ∧ Pertinenza(I, AmbCog(Dest)) ≥ θ]
dove θ è una soglia di pertinenza minima accettabile

(b) ¬ ∃ S' [C poteva produrre S' ∧ Pertinenza(S', AmbCog(Dest)) > Pertinenza(S, AmbCog(Dest))]
sotto vincoli di abilità e preferenze di C

Questa presunzione di pertinenza ottimale guida l'interpretazione attraverso quella che Sperber e Wilson chiamano la **Strategia di Comprensione Basata sulla Pertinenza** (Relevance-Guided Comprehension Strategy):

Passo 1: Segui un percorso di minimo sforzo nell'elaborazione dell'enunciato:

- Disambigua termini nell'ordine di accessibilità

- Risolvi riferimenti nell'ordine di accessibilità
- Arricchisci pragmaticamente nell'ordine di accessibilità
- Costruisci implicazioni contestuali nell'ordine di accessibilità

Passo 2: Arresta l'elaborazione quando: Le aspettative di pertinenza sono soddisfatte, ovvero quando hai derivato un'interpretazione che: (i) Produce effetti cognitivi sufficienti (ii) Richiede sforzo ragionevole (iii) È consistente con la presunzione di pertinenza ottimale

Questa strategia predice che gli ascoltatori accettano la **prima interpretazione che soddisfa le loro aspettative di pertinenza**, senza continuare a cercare interpretazioni potenzialmente più pertinenti. Questo cattura l'efficienza cognitiva della comprensione ordinaria: non investighiamo esaustivamente tutte le interpretazioni possibili ma procediamo euristicamente, arrestando la ricerca appena un'interpretazione adeguata è trovata.

Formalmente:

Interpretazione_Accettata(S, C) =
  Prima_Interpretazione_I tale che:
    I ∈ Spazio_Interpretazioni(S) ordinato per accessibilità
    ∧
    Pertinenza(I, C) ≥ Soglia_Aspettata(S, C)
    ∧
    Consistente_con_Presunzione_Ottimalità(I, S, C)

Esplicature e Implicature nella Teoria della Pertinenza

La Relevance Theory introduce una distinzione tecnica raffinata tra due livelli di comunicazione esplicita e implicita, diversa dalla distinzione griceana tradizionale tra "ciò che è detto" e "ciò che è implicato":

Esplicatura (Explicature): Una assunzione comunicata esplicitamente attraverso l'enunciato, costruita mediante sviluppo logico-pragmatico della forma logica codificata linguisticamente. L'esplicatura è una proposizione completamente articolata che risulta da:

- Decodifica del contenuto semantico
- Disambiguazione
- Risoluzione di riferimenti
- Arricchimento pragmatico
- Completamento di forme logiche sottodeterminate

Implicatura (Implicature): Una assunzione comunicata implicitamente, dedotta dal contenuto esplicito più il contesto, ma non articolata esplicitamente nell'enunciato

La distinzione cruciale è che l'esplicatura è uno **sviluppo** della forma logica codificata (contiene elementi semanticamente codificati dall'enunciato), mentre le implicature sono interamente **aggiunte** inferite pragmaticamente.

Consideriamo un esempio concreto:

(1) Enunciato: "Non ho mangiato" Contesto: Detto in risposta a un invito a cena, alla sera

Analisi della Teoria della Pertinenza:

Forma Logica Codificata:

$\exists e\ \exists t$ [Mangiare(e) $\wedge$ Agente(e, Io*) $\wedge$ Tempo(e, t) $\wedge$ t < Adesso* $\wedge$ ¬Passato(e)]

Dove * indica elementi indicali non saturati

Processi di Arricchimento per Derivare Esplicature:

a) Saturazione di indicali:

Io* → parlante = Giovanni
Adesso* → tempo di enunciazione = 19:00, 15 gennaio 2026

b) Risoluzione di sottodeterminazione temporale:

t < Adesso → t $\in$ {tempi recenti pertinenti}

c) Arricchimento pragmatico guidato da pertinenza: Dato il contesto (invito a cena serale), la restrizione temporale più pertinente è:

t $\in$ Questo_Giorno $\wedge$ t $\in$ Ora_Pranzo $\cup$ Tardo_Pomeriggio

d) Arricchimento del predicato: "Mangiare" → "Mangiare un pasto sostanziale" (non soltanto uno snack)

Esplicatura Finale:

Giovanni non ha mangiato un pasto sostanziale oggi durante le ore di pranzo o tardo pomeriggio

Questa è la proposizione esplicitamente comunicata, anche se gran parte del suo contenuto (restrizione temporale a oggi, specificazione di "pasto sostanziale", eccetera) non è semanticamente codificata ma pragmaticamente arricchita.

Implicature Derivate dall'Esplicatura:

Dato il contesto (invito a cena) e l'explicature, l'ascoltatore può derivare implicature attraverso ragionamento inferenziale:

Implicatura 1: "Giovanni ha fame"

Derivazione:

(i) Giovanni non ha mangiato sostanzialmente oggi [esplicatura]
(ii) Assunzione contestuale: Chi non mangia per diverse ore diventa affamato
(iii) Implicazione: Giovanni è probabilmente affamato [implicatura]

Implicatura 2: "Giovanni accetta l'invito a cena"

Derivazione:
(i) Giovanni non ha mangiato sostanzialmente oggi [esplicatura]
(ii) L'ascoltatore ha invitato Giovanni a cena
(iii) Assumendo cooperatività: Giovanni fornisce informazione pertinente all'invito
(iv) L'informazione che Giovanni non ha mangiato è pertinente soltanto se influenza la decisione
(v) Influenza favorevolmente (avendo fame, Giovanni vorrebbe cenare)
(vi) Implicazione: Giovanni intende accettare l'invito [implicatura]

Questa analisi mostra come l'explicature, pur contenendo contenuto semanticamente codificato, è sostanzialmente arricchita pragmaticamente, e come le implicature sono derivate inferenzialmente dall'esplicatura più il contesto attraverso ragionamento guidato da pertinenza.

Arricchimento Libero versus Saturazione Obbligatoria

Un'ulteriore distinzione teorica importante nella Relevance Theory concerne i tipi di processi pragmatici coinvolti nella costruzione dell'explicature:

Saturazione (Saturation): Riempimento obbligatorio di slots sottodeterminati semanticamente che devono essere riempiti per ottenere una proposizione completa valutabile per verità

Esempio: Indicali come "io", "qui", "ora" richiedono saturazione obbligatoria

"Io sono qui ora"
→ Richiede saturazione di: chi è "io", dove è "qui", quando è "ora"
Senza saturazione, non otteniamo una proposizione valutabile

Arricchimento Libero (Free Enrichment): Aggiunta opzionale di contenuto pragmatico che va oltre ciò che è semanticamente richiesto, guidata da considerazioni di pertinenza

Esempio: "Ho fatto colazione" → arricchimento temporale "questa mattina"

La forma logica è già una proposizione completa senza arricchimento
"Ho fatto colazione in qualche tempo passato" è valutabile per verità
Ma l'arricchimento a "questa mattina" lo rende più pertinente

Robyn Carston, una delle principali teoriche della Relevance Theory, argomenta che questa distinzione è in realtà sfumata e che molti casi apparentemente di saturazione obbligatoria sono meglio analizzati come arricchimento libero guidato da pertinenza. Per esempio:

(2) "È pronto"

Sembra richiedere saturazione di "esso" (cosa è pronto?) e possibilmente di "pronto per cosa?". Ma Carston sostiene che "pronto" non ha necessariamente un argomento sintattico sottostante che richiede saturazione. Invece, l'enunciato è semanticamente completo ma radicalmente sottodeterminato, e l'interpretazione pragmatica arricchisce liberamente specificando cosa è pronto e per cosa.

Analogamente: (3) "Tutti hanno superato l'esame"

"Tutti" sembra semanticamente non-saturato (tutti chi?). Ma Carston argumenta che "tutti" è semanticamente un quantificatore assoluto ($\forall x$) e la restrizione di dominio ("tutti gli studenti del corso", "tutti i candidati", eccetera) è pragmaticamente arricchita, non semanticamente richiesta.

Formalmente, possiamo distinguere i processi come:

Saturazione: Processo P è saturazione sse:

(i) La forma logica L contiene un elemento sottodeterminato σ
(ii) σ è una variabile libera o slot sintattico
(iii) L non è una proposizione completa finché σ non è assegnato un valore
(iv) P assegna un valore a σ basandosi su contesto

Arricchimento Libero: Processo P è arricchimento libero sse:

(i) La forma logica L è già una proposizione completa
(ii) P aggiunge contenuto addizionale C non richiesto per completezza proposizionale
(iii) L + C è più pertinente di L da solo nel contesto
(iv) La pertinenza extra giustifica lo sforzo addizionale di elaborare C

La questione empirica e teorica di dove tracciare il confine tra questi processi rimane dibattuta, con implicazioni per l'architettura della semantica-pragmatica interface.

Modalità Procedurale versus Concettuale del Significato

Un'innovazione teorica distintiva della Relevance Theory è la distinzione tra due tipi fondamentalmente diversi di contenuto semantico codificato dalle espressioni linguistiche:

Significato Concettuale (Conceptual Meaning): Espressioni che codificano concetti, ovvero rappresentazioni mentali che possono essere usate in pensiero e ragionamento proposizionale. I significati concettuali contribuiscono direttamente al contenuto proposizionale dell'enunciato.

Esempi: Nomi, verbi, aggettivi, avverbi

"gatto" → codifica il concetto GATTO
"correre" → codifica il concetto CORRERE
"velocemente" → codifica il concetto VELOCEMENTE

Significato Procedurale (Procedural Meaning): Espressioni che codificano istruzioni procedurali su come processare l'enunciato, ovvero vincoli su come manipolare rappresentazioni concettuali ma senza contribuire concetti al contenuto proposizionale.

Esempi: Marcatori discorsivi, pronomi, alcune particelle

"ma" → non codifica un concetto MA, ma piuttosto istruisce:
"Processa il seguente contenuto come contrastante con aspettative create dal contenuto precedente"

"quindi" → istruisce:
"Tratta il seguente contenuto come conclusione inferita dal contenuto precedente"

"anche" → istruisce:
"Aggiungi il seguente contenuto a un insieme di alternative già stabilite"

Consideriamo l'esempio classico che contrasta "ma" e "e":

(4a) "Il film era lungo ma interessante" (4b) "Il film era lungo e interessante"

(4a) e (4b) hanno lo stesso contenuto proposizionale concettuale:

$\text{Lungo(Film)} \wedge \text{Interessante(Film)}$

Ma differiscono nel significato procedurale:

- "e" semplicemente congiunge le due proprietà
- "ma" aggiunge un'istruzione procedurale: "Il secondo congiunto contrasta con aspettative create dal primo"

L'effetto comunicativo è diverso perché "ma" guida l'ascoltatore a derivare implicature addizionali:

Implicatura da (4a):

(i) I film lunghi sono tipicamente noiosi [assunzione contestuale]
(ii) "ma" segnala contrasto con aspettative
(iii) Il fatto che questo film lungo sia interessante contrasta con l'aspettativa di noia
(iv) Implicatura: Questo film è eccezionalmente interessante dato che è lungo

Formalmente, possiamo modellare significati procedurali come funzioni di ordine superiore che operano su rappresentazioni proposizionali:

⟦"ma"⟧ = λp . λq . λC . [p ∧ q ∧ Contrasta(q, Aspettative_da(p, C))]

dove:
p, q = contenuti proposizionali dei congiunti
C = contesto cognitivo
Aspettative_da(p, C) = insieme di assunzioni che seguirebbero normalmente da p in C
Contrasta(q, A) = q è inconsistente con o improbabile dato A

Altri esempi di codifica procedurale:

Pronomi: Non codificano concetti ma piuttosto istruzioni per recuperare referenti:

"lui" → Istruzione: "Accedi al referente maschile singolare più saliente nel contesto"

Tempi Verbali: Codificano istruzioni temporali procedurali oltre a contenuto concettuale:

Passato Prossimo Italiano:
- Contenuto concettuale: evento precede tempo di enunciazione
- Contenuto procedurale: "Tratta l'evento come rilevante per il presente / recente"

Passato Remoto:
- Contenuto concettuale: evento precede tempo di enunciazione
- Contenuto procedurale: "Tratta l'evento come disconnesso dal presente / remoto"

Particelle Modali: Codificano primariamente significato procedurale:

"davvero" → Istruzione: "Rafforza la forza assertiva dell'enunciato"
"forse" → Istruzione: "Indebolisci l'impegno epistemico all'enunciato"

Questa distinzione concettuale/procedurale ha implicazioni per l'architettura cognitiva del linguaggio: suggerisce che non tutto il significato linguistico consiste in contenuto proposizionale manipolabile nel ragionamento. Parte del significato consiste in istruzioni meta-rappresentazionali su come processare e integrare contenuto proposizionale.

La Semantica Lessicale e la Struttura dei Concetti Lessicalmente Codificati

Il Problema della Decomposizione Lessicale: Primitivi Semantici e Struttura Interna

Una questione centrale nella semantica lessicale riguarda se i significati delle parole sono entità atomiche non analizzabili ulteriormente, oppure se possiedono una struttura interna complessa decomponibile in componenti semantici più primitivi. Questa questione ha profonde ramificazioni per la teoria del significato, l'acquisizione lessicale, e l'architettura cognitiva del sistema concettuale.

La **teoria decomposizionale** (decompositional theory) sostiene che i significati lessicali possono e devono essere analizzati in termini di configurazioni strutturate di primitivi semantici. Questa tradizione risale all'analisi semantica componenziale sviluppata nell'antropologia linguistica da studiosi come Floyd Lounsbury e Ward Goodenough negli anni cinquanta, e successivamente adottata nella linguistica generativa attraverso il lavoro di Jerrold Katz, Jerry Fodor, e Ray Jackendoff.

Consideriamo l'analisi classica dei termini di parentela. In molte lingue, i termini di parentela possono essere analizzati come combinazioni di un numero ristretto di componenti semantici binari:

Componenti Primitivi per Termini di Parentela:

- [±MASCHIO]: distingue genere
- [±GENERAZIONE ASCENDENTE]: distingue generazioni superiori (genitori, nonni)
- [±GENERAZIONE DISCENDENTE]: distingue generazioni inferiori (figli, nipoti)
- [±LINEARITÀ]: distingue parenti diretti (genitori, figli) da collaterali (zii, cugini)
- [±CONSANGUINEITÀ]: distingue parenti di sangue da acquisiti (per matrimonio)

Utilizzando questi primitivi, possiamo decomporre:

"padre" = [+MASCHIO, +GENERAZIONE_ASCENDENTE, +LINEARE, +CONSANGUINEO]
"madre" = [-MASCHIO, +GENERAZIONE_ASCENDENTE, +LINEARE, +CONSANGUINEO]
"zio" = [+MASCHIO, +GENERAZIONE_ASCENDENTE, -LINEARE, +CONSANGUINEO]
"suocero" = [+MASCHIO, +GENERAZIONE_ASCENDENTE, +LINEARE, -CONSANGUINEO]
"figlio" = [+MASCHIO, +GENERAZIONE_DISCENDENTE, +LINEARE, +CONSANGUINEO]
"nipote" (figlio del figlio) = [±MASCHIO, +GENERAZIONE_DISCENDENTE_2, +LINEARE, +CONSANGUINEO]

Questa decomposizione spiega elegantemente relazioni di implicazione tra termini:

x è un padre → x è maschio
x è un padre → x ha almeno un figlio
x è uno zio di y → x è maschio $\wedge$ x è fratello di un genitore di y

Formalmente, possiamo rappresentare la decomposizione come una funzione che mappa parole a strutture di features:

Decomp: Lessema → P(Features)

dove Features = {[±F_1], [±F_2], ..., [±F□]}

Per esempio:
Decomp("padre") = {[+MASCHIO], [+GEN_ASC], [+LINEARE], [+CONSANG]}

Le relazioni semantiche tra lessemi possono essere definite in termini di sovrapposizioni e inclusioni di features:

Sinonimia (approssimativa):

Sinonimi(L_1, L_2) sse Decomp(L_1) = Decomp(L_2)

Antonimia:

Antonimi(L_1, L_2) sse Decomp(L_1) e Decomp(L_2) differiscono soltanto nel valore di un feature binario

Esempio: "padre" vs "madre" differiscono solo in [±MASCHIO]

Iponimia (relazione genere-specie):

Iponimo(L_1, L_2) sse Decomp(L_1) ⊃ Decomp(L_2)

Esempio: "stallone" è iponimo di "cavallo"

Decomp("cavallo") = {[+ANIMALE], [+EQUINO]}
Decomp("stallone") = {[+ANIMALE], [+EQUINO], [+MASCHIO], [+ADULTO]}

Questa teoria ha esteso l'analisi componenziale ben oltre i termini di parentela. Jackendoff, nel suo programma di **Semantica Concettuale** (Conceptual Semantics), propone decomposizioni elaborate per verbi di movimento, possesso, e cambiamento di stato.

Consideriamo i verbi di causazione e cambiamento:

"uccidere" = CAUSARE(x, DIVENTARE(y, MORTO))
"aprire" = CAUSARE(x, DIVENTARE(y, APERTO))
"rompere" = CAUSARE(x, DIVENTARE(y, ROTTO))

Dove CAUSARE, DIVENTARE, MORTO, APERTO, ROTTO sono primitivi semantici nella metalingua della decomposizione.

Questa decomposizione cattura generalizzazioni cross-lessicali. Per esempio, tutti i verbi causativi transitivi condividono la struttura CAUSARE(x, ...), e tutti i verbi di cambiamento di stato incorporano DIVENTARE(y, STATO).

Le inferenze semantiche seguono dalla struttura decomposta:

Da: Giovanni ha ucciso il serpente
Inferenza valida: Il serpente è diventato morto
Inferenza valida: Giovanni ha causato che il serpente morisse
Inferenza valida: Il serpente è morto

Da: Maria ha aperto la porta
Inferenza valida: La porta è diventata aperta
Inferenza valida: La porta è aperta (ora)

Primitivi Semantici Universali: Il Programma NSM di Wierzbicka

Anna Wierzbicka, nel suo ambizioso programma di ricerca durato decenni, ha cercato di identificare un insieme finito di **primitivi semantici universali** (semantic primes) che sono innati, universalmente presenti in tutte le lingue umane, e non decomponibili ulteriormente. Secondo la teoria del Metalinguaggio Semantico Naturale (Natural Semantic Metalanguage, NSM), questi primitivi costituiscono gli atomi irriducibili del significato umano, e tutti i concetti lessicali in tutte le lingue possono essere decomposti in configurazioni di questi primitivi.

Wierzbicka e collaboratori, attraverso ricerca tipologica estensiva su centinaia di lingue diverse, propongono un inventario di circa sessantacinque primitivi semantici universali, organizzati in categorie:

SOSTANTIVI (13 primitivi):

- IO, TU, QUALCUNO, QUALCOSA/COSA, GENTE, CORPO
- Primitivi relazionali: GENERE, PARTE, LATO

DETERMINATORI (4):

- QUESTO, LO STESSO, ALTRO/DIVERSO, UNO, DUE, ALCUNI, TUTTI, MOLTO/MOLTI, POCO/POCHI

VALUTATORI (2):

- BUONO, CATTIVO

DESCRITTORI (2):

- GRANDE, PICCOLO

QUANTIFICATORI MENTALI (3):

- PENSARE, SAPERE, VOLERE, SENTIRE, VEDERE, UDIRE

PREDICATI DI DISCORSO (3):

- DIRE, PAROLE, VERO

AZIONI E EVENTI (10):

- FARE, ACCADERE, MUOVERSI, TOCCARE
- ESSERCI/ESISTERE, AVERE, ESSERE (qualcosa/qualcuno)

VITA E MORTE (2):

- VIVERE, MORIRE

TEMPO (5):

- QUANDO/TEMPO, ORA, PRIMA, DOPO, MOLTO TEMPO, POCO TEMPO, PER QUALCHE TEMPO, MOMENTO

SPAZIO (8):

- DOVE/LUOGO, QUI, SOPRA, SOTTO, LONTANO, VICINO, LATO, DENTRO, TOCCARE

CONCETTI LOGICI (5):

- NON, FORSE, POTERE, A CAUSA DI, SE

INTENSIFICATORE, AUMENTO (2):

- MOLTO, PIÙ

SIMILARITÀ (1):

- COME/MODO

La tesi cruciale della NSM è che questi primitivi:

1. Sono **indefinibili**: non possono essere decomposti ulteriormente senza circolarità
2. Sono **universali**: hanno esponenti lessicali in ogni lingua naturale
3. Sono **auto-esplicativi**: i loro significati sono direttamente comprensibili senza necessità di definizione

Usando questi primitivi, Wierzbicka propone esplicazioni (explications) di concetti lessicali complessi attraverso parafrasi strutturate. Per esempio:

Esplicazione di "amico":

X è amico di Y =
  X pensa questo di Y:
    "questa persona (Y) è qualcuno come me
     io voglio fare cose buone per questa persona
     io penso che questa persona vuole fare cose buone per me
     io mi sento qualcosa di buono quando sono con questa persona"
  Y pensa lo stesso di X
  a causa di questo, X e Y fanno cose insieme qualche volta

Esplicazione di "pietà":

X sente pietà per Y =
  X pensa questo di Y:
    "qualcosa di cattivo è accaduto a questa persona
     questa persona si sente qualcosa di cattivo a causa di questo
     io non voglio che questa persona si senta qualcosa di cattivo"
  quando X pensa questo, X si sente qualcosa
  X si sente così perché X pensa questo

Esplicazione di "vergogna":

X si sente vergogna =
  X pensa questo:
    "le persone sanno qualcosa di cattivo di me
     a causa di questo, queste persone pensano qualcosa di cattivo di me
     io non voglio che le persone pensino questo di me
     io non posso fare niente ora"
  quando X pensa questo, X si sente qualcosa di cattivo

Queste esplicazioni utilizzano soltanto i primitivi identificati dalla NSM, rendendo i significati trasparenti e cross-linguisticamente comparabili.

Formalmente, possiamo concepire le esplicazioni NSM come formule in un linguaggio formale L_NSM:

L_NSM ::= Primitivo | Primitivo(L_NSM) | L_NSM $\wedge$ L_NSM | L_NSM $\rightarrow$ L_NSM | $\forall$x L_NSM | ...

dove Primitivo $\in$ {IO, TU, QUALCUNO, ..., COME}

Il significato di un lessema L in una lingua naturale è definito come:

Significato_NSM(L) = φ dove φ è una formula di L_NSM che esplica L

Due lessemi in lingue diverse sono **translation equivalents** sse hanno la stessa esplicazione NSM:

TranslEquiv(L_1, L_2) sse Significato_NSM(L_1) = Significato_NSM(L_2)

Questa teoria affronta critiche sostanziali. Primo, la selezione dei primitivi sembra in parte stipulativa: perché "buono" e "cattivo" sono primitivi mentre "bello" e "brutto" non lo sono? Secondo, molte esplicazioni NSM appaiono circuitose e perdono sfumature importanti dei

significati lessicali. Terzo, l'affermazione di universalità è controversa: alcuni concetti proposti come primitivi (per esempio "SAPERE") potrebbero non avere esponenti lessicali diretti in tutte le lingue.

Semantica Prototipica e Struttura Categoriale: La Rivoluzione di Rosch

Un approccio radicalmente diverso alla struttura lessicale emerge dalla psicologia cognitiva, particolarmente attraverso il lavoro pioneristico di Eleanor Rosch negli anni settanta. Rosch sfidò la concezione classica delle categorie, secondo cui le categorie sono definite da insiemi di condizioni necessarie e sufficienti (la concezione aristotelica che sottende la decomposizione componenziale), dimostrando empiricamente che molte categorie naturali hanno invece una **struttura prototipica**.

Secondo la **teoria del prototipo** (prototype theory):

1. Le categorie hanno **membri centrali** (prototipi) e **membri periferici** (esemplari meno tipici)
2. L'appartenenza alla categoria è una questione di **grado** piuttosto che tutto-o-niente
3. I confini categoriali sono **sfumati** (fuzzy) piuttosto che netti
4. I giudizi di tipicità sono sistematici e consistenti tra parlanti

Rosch condusse esperimenti dove chiedeva ai partecipanti di valutare la tipicità di vari esemplari per categorie come "uccello", "frutto", "mobile", "veicolo". I risultati mostrarono pattern consistenti:

Categoria: UCCELLO

Esemplari molto tipici (prototipici):
Pettirosso: tipicità 1,1 (su scala 1-7, dove 1 = molto tipico)
Passero: 1,2
Colomba: 1,4

Esemplari moderatamente tipici:
Aquila: 2,3
Pappagallo: 3,1
Falco: 3,3

Esemplari poco tipici (periferici):
Struzzo: 5,8
Pinguino: 6,2
Pipistrello: 6,9 (erroneamente categorizzato da alcuni come uccello)

Categoria: FRUTTO

Molto tipici: Mela (1,3), Arancia (1,4), Banana (1,5)
Moderatamente tipici: Fragola (2,3), Pesca (2,5), Uva (2,6)
Poco tipici: Oliva (5,1), Cocco (5,5), Zucca (6,2)

Questi effetti di tipicità si manifestano in diversi compiti cognitivi:

Tempo di Verifica: I soggetti verificano più rapidamente affermazioni categoriali per membri prototipici:

"Un pettirosso è un uccello" → tempo di reazione: 610 ms
"Un pinguino è un uccello" → tempo di reazione: 875 ms

Apprendimento Categoriale: Bambini apprendono prima le categorie attraverso esemplari prototipici, poi estendono a membri meno tipici.

Effetti di Priming: Membri prototipici fungono da migliori primes per la categoria:

Prime: "uccello" → Target: "pettirosso" (forte facilitazione)
Prime: "uccello" → Target: "pinguino" (facilitazione debole)

Rosch interpreta questi dati come evidenza che le categorie sono organizzate attorno a **prototipi**, che sono rappresentazioni mentali degli esemplari più tipici o ideali della categoria. La categorizzazione procede valutando la similarità di un nuovo esemplare al prototipo:

Grado_di_Appartenenza(x, Categoria_C) = Similarità(x, Prototipo_C)

Formalmente, possiamo modellare una categoria prototipica C come una funzione di appartenenza fuzzy:

μ_C: Dominio → [0, 1]

dove $\mu_C(x)$ è il grado in cui x appartiene a C

Il prototipo è l'elemento con grado massimo:

Prototipo_C = argmax[x] $\mu_C(x)$

La similarità può essere definita in termini di features condivise:

Similarità(x, y) = |Features(x) ∩ Features(y)| / |Features(x) ∪ Features(y)|

Questa concezione contrasta nettamente con la teoria classica delle categorie, dove l'appartenenza è binaria:

Teoria Classica: x ∈ C sse x soddisfa tutte le condizioni necessarie e sufficienti per C
$\mu_C(x) \in \{0, 1\}$

Teoria Prototipica: x appartiene a C in grado μ_C(x) ∈ [0, 1]
basato su similarità al prototipo

Livelli di Categorizzazione: Livello di Base, Superordinato, Subordinato

Un'altra scoperta importante di Rosch riguarda i **livelli di categorizzazione**. Le categorie sono organizzate in gerarchie tassonomiche con tre livelli principali:

Livello Superordinato (Superordinate Level): Categorie molto generali e astratte Esempi: MOBILE, ANIMALE, VEICOLO, STRUMENTO

Livello di Base (Basic Level): Livello intermedio che è cognitivamente privilegiato Esempi: SEDIA, CANE, AUTO, MARTELLO

Livello Subordinato (Subordinate Level): Categorie molto specifiche Esempi: SEDIA_A_DONDOLO, GOLDEN_RETRIEVER, FERRARI_TESTAROSSA, MARTELLO_DA_CARPENTIERE

Rosch dimostrò che il **livello di base** ha uno status speciale nella cognizione umana:

1. **Primo appreso dai bambini**: I bambini apprendono termini di livello di base ("cane", "sedia") prima di termini superordinati ("animale", "mobile") o subordinati ("collie", "poltrona")

2. **Usato preferenzialmente nella denominazione spontanea**: Quando mostrati un oggetto e chiesti di nominarlo, gli adulti preferenzialmente usano il termine di livello di base

3. **Massima distintività percettiva**: Gli oggetti in categorie di livello di base diverse hanno forme percettivamente molto distinte, mentre oggetti in categorie subordinate sono percettivamente simili

4. **Massima informazione al minimo costo cognitivo**: Il livello di base massimizza il rapporto tra informazione guadagnata (distintività) e sforzo cognitivo (numero di distinzioni necessarie)

Formalmente, possiamo caratterizzare i livelli in termini di bilancio informazione/costo:

Livello_Superordinato:
- Costo cognitivo: Basso (poche distinzioni necessarie)
- Informazione: Bassa (alta astrazione, pochi dettagli)
- Utilità: (Informazione / Costo) = Bassa / Basso = Media

Livello_di_Base:
- Costo cognitivo: Medio
- Informazione: Alta (buona specificità senza dettagli eccessivi)
- Utilità: Alta / Medio = MASSIMA

Livello_Subordinato:
- Costo cognitivo: Alto (molte distinzioni sottili richieste)
- Informazione: Molto alta
- Utilità: Molto Alta / Alto = Media

Il livello di base rappresenta un **ottimale cognitivo** dove l'informazione guadagnata è massimizzata relativamente allo sforzo richiesto.

Inoltre, il livello di base corrisponde al livello dove:

Massima somiglianza intra-categoria: Membri della stessa categoria di livello di base sono molto simili tra loro

Similarità_Intra(SEDIA) > Similarità_Intra(MOBILE)
Tutte le sedie condividono molte features percettive

Massima differenza inter-categoria: Categorie di livello di base diverse sono molto dissimili

Dissimilarità(SEDIA, TAVOLO) > Dissimilarità(SEDIA_A_DONDOLO, SEDIA_DA_UFFICIO)

Massimo numero di features comuni: Membri di categoria di livello di base condividono molte features motorie e funzionali

Azioni_motorie(SEDIA) = {sedersi, appoggiarsi, sollevare, ...}
Queste azioni sono simili per tutte le sedie ma diverse per altri mobili

Teoria della Metafora Concettuale: Lakoff e Johnson

George Lakoff e Mark Johnson, nel loro libro rivoluzionario "Metaphors We Live By" del millenovecentottanta, proposero che la metafora non è semplicemente un ornamento retorico o una figura del discorso poetico, ma piuttosto è un meccanismo cognitivo fondamentale attraverso cui strutturiamo il nostro sistema concettuale. Secondo la **teoria della metafora concettuale** (conceptual metaphor theory), gran parte del nostro pensiero astratto è strutturato metaforicamente attraverso mappature sistematiche (mappings) da domini concreti e percettivamente basati (domini sorgente) a domini astratti (domini bersaglio).

Una metafora concettuale ha la forma generale:

DOMINIO_BERSAGLIO è DOMINIO_SORGENTE

dove elementi e relazioni nel dominio sorgente sono sistematicamente mappati su elementi e relazioni nel dominio bersaglio.

Esempio paradigmatico: **L'AMORE È UN VIAGGIO**

Questa metafora concettuale licenza un'intera famiglia di espressioni metaforiche linguistiche:

Espressioni linguistiche basate su AMORE-È-VIAGGIO:

"La nostra relazione è a un bivio"
  Mappatura: Decisione in relazione ↔ Bivio nel viaggio

"Questa relazione non sta andando da nessuna parte"
  Mappatura: Progresso relazionale ↔ Avanzamento verso destinazione

"Guardate quanto abbiamo fatto strada insieme"
  Mappatura: Sviluppo della relazione ↔ Distanza percorsa

"Abbiamo colpito una barriera"
  Mappatura: Ostacolo relazionale ↔ Ostacolo fisico nel viaggio

"È stata una relazione accidentata"
  Mappatura: Difficoltà relazionali ↔ Terreno difficile

"Siamo a un punto morto"
  Mappatura: Stallo relazionale ↔ Arresto del movimento

La mappatura sistematica è:

DOMINIO SORGENTE (VIAGGIO):
  - Viaggiatori → Amanti
  - Veicolo → Relazione amorosa
  - Destinazione → Scopi relazionali (matrimonio, convivenza, etc.)
  - Distanza percorsa → Progresso nella relazione
  - Ostacoli nel percorso → Difficoltà relazionali
  - Bivii → Scelte cruciali
  - Guida del veicolo → Controllo sulla relazione

Formalmente, una metafora concettuale M: S → T è una funzione parziale che mappa elementi del dominio sorgente S a elementi del dominio bersaglio T:

M: Ontologia(S) → Ontologia(T)

dove Ontologia(D) = {entità, relazioni, proprietà in dominio D}

Per la metafora AMORE-È-VIAGGIO:

M_viaggio-amore: Ontologia(VIAGGIO) → Ontologia(AMORE)

M(viaggiatore) = amante
M(veicolo) = relazione
M(destinazione) = scopo_relazionale
M(percorrere_distanza) = progredire_relazione
M(ostacolo) = difficoltà
...

Le metafore concettuali non sono arbitrarie ma sono **motivate** dalla nostra esperienza corporea e fisica. Lakoff e Johnson identificano tre tipi principali di motivazione:

1. Metafore Orientazionali: Basate su orientamenti spaziali fondamentali del corpo umano (su/giù, dentro/fuori, davanti/dietro, centrale/periferico)

FELICE È SU / TRISTE È GIÙ
"Mi sento su di morale" vs "Sono giù"
"Il morale è alto" vs "È caduto in depressione"
Motivazione: Postura corporea eretta associata a vitalità, curvata a debolezza

PIÙ È SU / MENO È GIÙ
"I prezzi stanno salendo"
"La temperatura è scesa"
Motivazione: Aggiungere più sostanza a un contenitore aumenta il livello

CONTROLLO È SU / ESSERE CONTROLLATO È GIÙ
"Domina la situazione" (è sopra)
"È caduto sotto il suo controllo"
Motivazione: Vincere fisicamente significa finire sopra l'avversario

2. Metafore Ontologiche: Concepiscono eventi astratti, emozioni, idee come entità fisiche, sostanze, o contenitori

LA MENTE È UN CONTENITORE
"Non riesco a mettere quell'idea nella mia testa"
"È pieno di rabbia repressa"
"Lasciami che ti riempia di idee"

LE TEORIE SONO EDIFICI
"La teoria ha fondamenta solide"
"L'argomento è crollato sotto critica"
"Dobbiamo costruire su questa base"

IL TEMPO È UNA RISORSA / DENARO
"Stai sprecando il mio tempo"
"Questo mi farà risparmiare ore"
"Investi tempo in questo progetto"

"Ho speso troppo tempo su questo"

3. Metafore Strutturali: Mappature ricche che strutturano un intero dominio astratto in termini di un altro dominio

L'ARGOMENTO È GUERRA
"Le tue affermazioni sono indifendibili"
"Ha attaccato ogni punto debole nel mio argomento"
"Le sue critiche hanno colpito nel segno"
"Ho demolito il suo argomento"
"Non ho mai vinto una discussione con lui"

LE IDEE SONO CIBO
"Quella è un'idea cruda"
"Lasciami che digerisca quella teoria"
"È un'idea difficile da ingoiare"
"Un'idea ben cotta"
"Stai rumando su questa idea"
"Il seminario è stato cibo per la mente"

Inferenze Metaforiche e Conseguenze Ontologiche

Crucialmente, le metafore concettuali non sono semplicemente modi di parlare, ma guidano il nostro ragionamento inferenziale. Le inferenze valide nel dominio sorgente vengono proiettate sul dominio bersaglio attraverso la mappatura metaforica:

Esempio: Inferenze attraverso TEORIE-SONO-EDIFICI

Fatto nel dominio sorgente (EDIFICI):
Se le fondamenta di un edificio sono deboli, l'edificio crollerà

Mappatura metaforica:
Fondamenta_edificio → Assunzioni_fondamentali_teoria
Crollo_edificio → Fallimento_teoria

Inferenza nel dominio bersaglio (TEORIE):
Se le assunzioni fondamentali di una teoria sono deboli, la teoria fallirà

Questa inferenza è sistematicamente sfruttata in argomentazione scientifica e filosofica.

Formalmente, il principio di **invarianza** (invariance principle) di Lakoff afferma che le mappature metaforiche preservano la struttura cognitiva-topologica del dominio sorgente:

Principio di Invarianza:
Se R(x, y) è una relazione strutturale nel dominio sorgente S,
e M: S → T è una metafora concettuale,

allora R'(M(x), M(y)) è una relazione corrispondente nel dominio bersaglio T

dove R' preserva la topologia cognitiva di R

Per esempio, nella metafora VITA-È-VIAGGIO:

Relazione sorgente: Essere_prima_nel_percorso(x, y) nel dominio VIAGGIO
Mappatura: Tappe_vita ↔ Punti_nel_viaggio
Relazione bersaglio: Essere_prima_cronologicamente(x, y) nel dominio VITA

L'ordinamento temporale nella vita è strutturato come ordinamento spaziale nel viaggio

Le metafore concettuali hanno anche **conseguenze ontologiche**: come concepiamo metaforicamente un dominio influenza quali entità postuliamo come esistenti in quel dominio.

Metafora: LA MENTE È UN CONTENITORE

Conseguenze ontologiche:
- La mente diventa concepita come avente confini (dentro/fuori)
- I pensieri diventano oggetti che possono essere "dentro" la mente
- Capacità mentale diventa "spazio" disponibile
- Emerge nozione di "contenuto mentale"

Espressioni risultanti:
"Togliti quell'idea dalla testa"
"Non entra nella mia testa"
"La mia mente è piena"

Lakoff argomenta che molti concetti filosofici tradizionali (sostanza, essenza, causalità, sé) sono parzialmente costituiti da sistemi metaforici concettuali, e che l'analisi filosofica deve riconoscere questo substrato metaforico.

Polisemia Sistematica e Estensioni Metaforiche

La teoria della metafora concettuale fornisce anche una spiegazione per pattern sistematici di **polisemia** (parole con molteplici significati correlati). Molti casi di polisemia riflettono estensioni metaforiche regolari:

Esempio: Verbo "vedere"

Significato letterale (percezione visiva):

"Vedo un albero davanti a me"

Estensioni metaforiche sistematiche basate su COMPRENDERE-È-VEDERE:

"Vedo cosa intendi" (comprendere)
"Vedi il punto dell'argomento?" (comprendere la rilevanza)
"Non vedo come questo possa funzionare" (comprendere la possibilità)
"Ah, ora lo vedo!" (momento di comprensione)

La mappatura metaforica COMPRENDERE-È-VEDERE è motivata esperienzialmente: la visione è il nostro senso dominante per acquisire informazione sul mondo, quindi serve come dominio sorgente naturale per concettualizzare acquisizione di conoscenza intellettuale.

Mappatura:
Vedere_oggetto → Comprendere_idea
Campo_visivo → Ambito_di_comprensione
Chiarezza_visiva → Chiarezza_intellettuale
Illuminazione → Spiegazione

Inferenze preservate:
Più luce → migliore visione ⇒ Più spiegazione → migliore comprensione
Ostacoli_visivi bloccano visione ⇒ Ostacoli_intellettuali bloccano comprensione

Altri pattern sistematici di polisemia metaforica:

"Caldo/Freddo" applicato a personalità:

Letterale: temperatura fisica
Metaforico: affettività sociale

"Una persona calda" = affettuosa, amichevole
"Una persona fredda" = distante, non emotiva
"Un'accoglienza calorosa"
"Un clima freddo in ufficio"

Basato su: AFFETTO-È-CALORE
Motivazione: Calore fisico associato a vicinanza corporea e cure parentali

"Alto/Basso" applicato a status:

Letterale: posizione verticale
Metaforico: gerarchia sociale

"Alta società"
"Caduto in basso"
"Scalare la gerarchia"
"Guardare dall'alto in basso qualcuno"

Basato su: STATUS-È-SU / BASSO-STATUS-È-GIÙ
Motivazione: Dominanza fisica coinvolge essere sopra fisicamente

Formalmente, possiamo modellare la polisemia sistematica come una rete di significati connessi da mappature metaforiche:

Rete_Polisemica(Parola_W) = ⟨{S_1, S_2, ..., S_n}, {M_{12}, M_{13}, ..., M_{in}}⟩

dove:
S_i = i-esimo senso di W
M_{in} = mappatura metaforica che connette S_i a S_n

Tipicamente:
S_1 = senso primario/letterale (dominio sorgente)
S_2, ..., S_n = sensi derivati/metaforici (domini bersaglio)

Per "vedere":

Rete_Polisemica("vedere") =
⟨{Percezione_visiva, Comprendere, Immaginare, Predire, ...},
{M_vedere-comprendere, M_vedere-immaginare, ...}⟩

dove ogni M è una metafora concettuale specifica

Questa struttura di rete contrasta con teorie che trattano sensi polisemici come entries lessicali completamente indipendenti. La teoria metaforica predice che i sensi sono sistematicamente correlati e che le correlazioni riflettono pattern cognitivi profondi.

Le Basi Neurali della Comprensione Linguistica: Architettura Funzionale e Substrati Cerebrali

Organizzazione Neuroanatomica del Sistema Linguistico: Modelli Classici e Revisioni Contemporanee

La neurolinguistica, ovvero lo studio delle basi neurali del linguaggio, ha una storia che risale al diciannovesimo secolo con le scoperte pionieristiche di Paul Broca e Carl Wernicke, che identificarono regioni cerebrali specifiche criticamente coinvolte nella produzione e comprensione linguistica. Il modello classico Broca-Wernicke, sebbene sostanzialmente revisionato dalla ricerca contemporanea, fornisce ancora un punto di partenza concettuale utile.

Modello Classico Broca-Wernicke (formulazione ottocentesca):

Il cervello contiene due regioni primarie specializzate per il linguaggio, entrambe tipicamente localizzate nell'emisfero sinistro:

Area di Broca (pars opercularis e pars triangularis del giro frontale inferiore, approssimativamente area di Brodmann quarantaquattro e quarantacinque):

- Funzione proposta: Produzione linguistica, particolarmente articolazione motoria del discorso e sintassi
- Evidenza: Lesioni in questa regione producono afasia di Broca (afasia espressiva o non-fluente)

Caratteristiche cliniche dell'afasia di Broca:

Produzione linguistica:
- Discorso non-fluente, telegrafico
- Agrammatismo (omissione di morfemi funzionali, flessioni)
- Difficoltà nell'articolazione (aprassia del discorso)
- Esempio: "Ragazzo... mangiare... mela" per "Il ragazzo sta mangiando una mela"

Comprensione:
- Relativamente preservata per frasi semplici
- Compromessa per frasi sintatticamente complesse
- Difficoltà con reversibilità semantica

Area di Wernicke (porzione posteriore del giro temporale superiore, approssimativamente area di Brodmann ventidue):

- Funzione proposta: Comprensione linguistica, particolarmente processamento semantico-lessicale
- Evidenza: Lesioni producono afasia di Wernicke (afasia recettiva o fluente)

Caratteristiche cliniche dell'afasia di Wernicke:

Produzione linguistica:
- Discorso fluente ma privo di contenuto informativo
- Parafasie semantiche (sostituzioni di parole: "forchetta" → "coltello")
- Neologismi (parole inventate)
- "Insalata di parole" (word salad) nei casi severi
- Esempio: "Sì, bene, ieri quando il tavolo ha fatto la cosa con il tempo, ho visto il gabbiano verde"

Comprensione:
- Severamente compromessa
- Incapacità di comprendere anche enunciati semplici
- Spesso inconsapevolezza del deficit (anosognosia)

Fascicolo Arcuato: Fascio di fibre di sostanza bianca che connette area di Broca e area di Wernicke

- Funzione proposta: Trasmissione di informazione tra regioni di produzione e comprensione
- Evidenza: Lesioni producono afasia di conduzione

Afasia di conduzione:

- Produzione relativamente fluente
- Comprensione relativamente preservata
- Deficit severo nella ripetizione
- Incapacità di correggere errori fonemici (parafasie fonemiche)

Il modello classico proponeva un flusso seriale di elaborazione:

COMPRENSIONE:
Input uditivo → Corteccia uditiva primaria (A1)
→ Area di Wernicke (decodifica fonologica e lessicale)
→ Fascicolo arcuato → Area di Broca (elaborazione sintattica)
→ Corteccia prefrontale (integrazione semantica)

PRODUZIONE:
Intenzione comunicativa → Corteccia prefrontale (pianificazione)
→ Area di Wernicke (selezione lessicale)
→ Fascicolo arcuato → Area di Broca (assemblaggio sintattico)
→ Corteccia motoria primaria (M1) → Output articolatorio

Formalmente, possiamo rappresentare questo modello come un grafo diretto:

$G_{classico} = \langle V, E \rangle$

$V = \{A1, Wernicke, Broca, FA, M1, PFC\}$
dove:
A1 = corteccia uditiva primaria
FA = fascicolo arcuato
M1 = corteccia motoria primaria
PFC = corteccia prefrontale

$E = \{(A1, Wernicke), (Wernicke, FA), (FA, Broca), (Broca, M1), (Wernicke, PFC), (PFC, Broca)\}$

Revisioni Contemporanee: Il Modello a Doppio Flusso e Reti Distribuiti

La ricerca neuroimaging contemporanea, utilizzando risonanza magnetica funzionale (fMRI), magnetoencefalografia (MEG), e stimolazione cerebrale, ha sostanzialmente revisionato e

arricchito il modello classico. Un framework influente contemporaneo è il **modello a doppio flusso** (dual-stream model) proposto da Gregory Hickok e David Poeppel.

Modello a Doppio Flusso (Hickok & Poeppel, duemilaquattro-duemilaotto):

Il processamento linguistico coinvolge due flussi di elaborazione paralleli e parzialmente indipendenti:

Flusso Ventrale (Ventral Stream): "Via del Cosa" (What pathway)

- Traiettoria anatomica: Corteccia uditiva (giro temporale superiore) → giro temporale medio → giro temporale inferiore → polo temporale
- Funzione: Mappatura suono-a-significato, comprensione semantica, riconoscimento lessicale
- Bilaterale (entrambi gli emisferi contribuiscono)

Processamento nel flusso ventrale:
Input acustico → Analisi spettro-temporale (STG bilaterale)
→ Rappresentazioni fonologiche (STG posteriore)
→ Interfaccia fonologia-semantica (MTG/ITG)
→ Rappresentazioni concettuali (polo temporale, corteccia prefrontale ventrale)

Flusso Dorsale (Dorsal Stream): "Via del Come" (How pathway)

- Traiettoria anatomica: Corteccia uditiva → corteccia parietale inferiore (giro sopramarginale, giro angolare) → corteccia premotoria/frontale inferiore (area di Broca)
- Funzione: Mappatura suono-a-articolazione, elaborazione fonologica di working memory, interfaccia sensori-motoria
- Fortemente lateralizzata a sinistra

Processamento nel flusso dorsale:
Rappresentazioni fonologiche → Giro sopramarginale (buffer fonologico)
→ Area Spt (area Silviana parietotemporale)
→ Corteccia premotoria (programmi motori articolatori)
→ Area di Broca (controllo articolatorio fine)

Questo modello spiega diversi pattern di dissociazione neurologica:

Dissociazione 1: Comprensione preservata con ripetizione compromessa

Lesione: Flusso dorsale (particolarmente giro sopramarginale, area Spt)
Risultato: Flusso ventrale intatto → comprensione OK
Flusso dorsale danneggiato → ripetizione compromessa
Sindrome clinica: Afasia di conduzione

Dissociazione 2: Comprensione compromessa con ripetizione preservata

Lesione: Flusso ventrale (giro temporale medio/inferiore)
Risultato: Flusso dorsale intatto → ripetizione OK (via mappatura acustica-articolatoria)
Flusso ventrale danneggiato → comprensione compromessa
Sindrome clinica: Afasia transcorticale sensoriale

Formalmente, il modello a doppio flusso è:

$G_{doppio\text{-}flusso} = \langle V, E_{ventrale} \cup E_{dorsale} \rangle$

$V = \{A1, STG_{post}, MTG, ITG, TP, SMG, AG, Spt, PM, IFG\}$

$E_{ventrale} = \{(A1, STG_{post}), (STG_{post}, MTG), (MTG, ITG), (ITG, TP)\}$
Funzione: λ input . Semantica(Lessico(Fonologia(input)))

$E_{dorsale} = \{(A1, STG_{post}), (STG_{post}, SMG), (SMG, Spt), (Spt, PM), (PM, IFG)\}$
Funzione: λ input . Articolazione(MotorProgram(Fonologia(input)))

Substrati Neurali del Processamento Sintattico: Evidenze da Neuroimaging e Studi Lesionali

Una questione centrale in neurolinguistica riguarda se esistono regioni cerebrali specificamente dedicate al processamento sintattico, distinte da regioni per processamento semantico. La ricerca ha prodotto evidenze convergenti per una **rete sintattica** distribuita ma parzialmente localizzata.

Regioni critiche per sintassi:

1. Giro Frontale Inferiore Sinistro (Left Inferior Frontal Gyrus, LIFG), particolarmente pars opercularis (BA44):

Evidenze fMRI:

Contrasto: Frasi sintatticamente complesse vs frasi semplici
Attivazione: LIFG (BA44) mostra incremento robusto

Esempio contrasto:
Complessa: "Il gatto che il cane che il ragazzo vede insegue corre"
Semplice: "Il ragazzo vede il cane e il gatto corre"

Risultato: BA44 ↑↑ per frasi complesse

Evidenze da stimolazione corticale diretta:

Procedura: Durante chirurgia cerebrale sveglia, elettrodi stimolano BA44
Compito: Produrre frasi grammaticali
Risultato: Stimolazione → agrammatismo transitorio

"Il ragazzo mangia la mela" → "Ragazzo... mela... mangiare"

2. Giro Temporale Superiore Posteriore (Posterior Superior Temporal Gyrus, pSTG), bilaterale ma con dominanza sinistra:

Evidenze MEG:

Paradigma: Violazioni sintattiche (phrase structure violations)
Stimoli: "Il gatto sui tetto dorme" vs "Il gatto sul tetto dorme"

Risposta neurale:
Latenza: 300-500 ms post-parola critica
Localizzazione: pSTG sinistro mostra early anterior negativity (ELAN)
Interpretazione: Rilevamento precoce automatico di violazioni categoriali

3. Giro Temporale Medio Sinistro (Left Middle Temporal Gyrus, LMTG):

Evidenze da afasia:

Lesione: LMTG anteriore
Deficit: Comprensione di frasi con strutture argomentali complesse

Test: Sentence-Picture Matching con frasi reversibili
"Il cane insegue il gatto" vs "Il gatto insegue il cane"

Pazienti con lesione LMTG: performance a livello di caso (50%)
Interpretazione: Ruolo critico nella mappatura sintassi-semantica (thematic role assignment)

4. Giro Angolare Sinistro (Left Angular Gyrus, LAG):

Evidenze fMRI:

Contrasto: Frasi con movimento sintattico vs frasi canoniche
Movimento: "Quale libro ha Maria letto __?"
Canoniche: "Maria ha letto quale libro?"

Risultato: LAG mostra attivazione per tracce sintattiche (gap-filling)
Interpretazione: Mantiene dipendenze a lunga distanza in memoria sintattica

Integrazione: La rete sintattica distribuita

La sintassi non è localizzata in un singolo "centro sintattico" ma emerge dall'interazione di una rete distribuita:

Rete_Sintattica = {LIFG(BA44), LIFG(BA45), pSTG, LMTG, LAG, Giro_Sopramarginale}

Divisione funzionale:
LIFG(BA44): Combinazione sintattica, merge operations
LIFG(BA45): Selezione lessicale con vincoli sintattici
pSTG: Costruzione di struttura sintattica online
LMTG: Integrazione sintassi-semantica
LAG: Memoria di dipendenze sintattiche
Giro Sopramarginale: Buffer fonologico per sintassi

Formalmente, possiamo modellare l'elaborazione sintattica come computazione distribuita:

SyntaxComputation(Input_Sentence) =
Σ[Region ∈ Rete_Sintattica] w_Region · f_Region(Input_Sentence)

dove:
f_Region = funzione computazionale specifica della regione
w_Region = peso di contributo (derivato da connettività effettiva)

Dinamica Temporale del Processamento Linguistico: Evidenze da Potenziali Evento-Correlati

L'elettroencefalografia (EEG) e la magnetoencefalografia (MEG) forniscono risoluzione temporale millisecondi per millisecondi del processamento linguistico, rivelando la dinamica temporale fine degli eventi neurali sottostanti la comprensione.

Componenti ERP (Event-Related Potentials) linguistici principali:

N100: Negatività a circa cento millisecondi

Latenza: 80-120 ms
Distribuzione: Corteccia uditiva bilaterale (elettrodi temporali)
Funzione: Risposta precoce a stimoli acustici
Sensibilità: Intensità acustica, cambiamenti fisici nel segnale
Interpretazione: Elaborazione percettiva pre-lessicale

N170: Negatività a circa centosettanta millisecondi

Latenza: 150-200 ms
Distribuzione: Occipito-temporale
Funzione: Riconoscimento visivo di parole scritte
Sensibilità: Expertise ortografica (maggiore per lettere vs simboli)
Correlato funzionale: Attivazione della Visual Word Form Area (VWFA)

N250: Negatività a circa duecentocinquanta millisecondi

Latenza: 200-300 ms

Funzione: Accesso lessicale precoce, rappresentazioni fonologiche
Sensibilità: Frequenza lessicale (ridotta per parole frequenti)
Ripetizione lessicale (priming)
Interpretazione: Prime interazioni lessico-fonetiche

N400: Negatività a circa quattrocento millisecondi (componente più studiata)

Latenza: 300-500 ms
Distribuzione: Centro-parietale, predominanza posteriore
Funzione: Integrazione semantica, accesso memoria semantica

Modulazioni N400:

1. Anomalia semantica:
"Bevo il caffè con zucchero e calzini"
→ "calzini" evoca N400 massiva

2. Plausibilità:
"Il bambino dorme nel letto" (N400 piccola)
"Il bambino dorme nel frigo" (N400 grande)

3. Congruenza contestuale:
Contesto: "Aveva fame dopo la corsa"
Parola finale: "mangiò una pizza" (N400 ridotta - congruente)
"mangiò una scarpa" (N400 aumentata - incongruente)

4. Frequenza lessicale:
Parole rare → N400 maggiore
Parole frequenti → N400 ridotta

5. Ripetizione:
Prima presentazione: N400 grande
Seconda presentazione: N400 ridotta (priming)

6. Relazioni semantiche:
Prime: "dottore" → Target: "infermiere" (N400 ridotta)
Prime: "tavolo" → Target: "infermiere" (N400 normale)

L'ampiezza N400 è inversamente correlata alla facilità di integrazione semantica:

$$\text{Ampiezza\_N400} \propto 1 / \text{Facilità\_Integrazione\_Semantica}$$

$$\text{Facilità\_Integrazione} = f(\text{Plausibilità}, \text{Relatedness\_Semantica}, \text{Aspettativa\_Contestuale})$$

P600 / SPS (Syntactic Positive Shift): Positività a circa seicento millisecondi

Latenza: 500-900 ms
Distribuzione: Centro-parietale, posteriore
Funzione: Rianalisi sintattica, riparazione di errori sintattici, integrazione strutturale complessa

Modulazioni P600:

1. Violazioni sintattiche:
   Grammaticale: "Il ragazzo che la ragazza ha baciato era alto"
   Violazione: "Il ragazzo che la ragazza ha baciato erano alti"
   → "erano" evoca P600

2. Garden-path (frasi ambigue):
   "Mentre Maria vestiva il bambino giocava nel parco"
   → "giocava" richiede rianalisi (Maria non vestiva il bambino, ma si vestiva lei stessa)
   → P600 su "giocava"

3. Complessità sintattica:
   Frasi con multiple embeddings evocano P600 sostenuta

4. Thematic role violations:
   "La bistecca ha mangiato..."
   → Violazione selezione argomentale → P600

ELAN (Early Left Anterior Negativity): Negatività precoce anteriore sinistra

Latenza: 100-300 ms
Distribuzione: Frontale sinistra
Funzione: Rilevamento automatico precoce di violazioni categoriali sintattiche

Esempio:
  Grammaticale: "Der Freund wurde..." (L'amico è stato...)
  Violazione: "*Der Freund wurde-st..." (categoria verbale errata)
  → ELAN a 150 ms su morfema scorretto

LAN (Left Anterior Negativity): Negatività anteriore sinistra

Latenza: 300-500 ms
Distribuzione: Anteriore sinistra (frontotemporale)
Funzione: Elaborazione morfosinttatica, accordo, working memory sintattica

Modulazioni:

1. Violazioni di accordo:
   "Il ragazzo mangiano" → LAN su "mangiano"

2. Violazioni di caso:
Tedesco: "*Den Vater schläft" (accusativo invece di nominativo)
→ LAN

3. Processi di working memory sintattico:
Mantenimento di dipendenze a distanza

Time Course dell'Elaborazione: Integrazione Multi-Livello

Combinando evidenze da ERP, possiamo ricostruire il corso temporale dell'elaborazione linguistica con precisione millisecondi:

Timeline del Processamento di una Parola:

0-100 ms: Elaborazione acustica/percettiva precoce
- N100: Risposta a onset acustico
- Corteccia uditiva primaria bilaterale
- Estrazione features acustiche di basso livello

100-200 ms: Pre-attivazione lessicale
- Attivazione coorte lessicale basata su input fonologico parziale
- Competizione tra candidati lessicali
- ELAN: Rilevamento categoriale (se violazione presente)

200-300 ms: Accesso lessicale
- N250: Selezione lessicale
- Attivazione rappresentazioni semantiche associate
- Effetti di frequenza lessicale emergono

300-400 ms: Integrazione semantica iniziale
- N400 inizia
- Integrazione della parola con contesto semantico precedente
- Computazione plausibilità e coerenza

400-500 ms: Integrazione semantica completa
- Picco N400
- Risoluzione ambiguità semantiche
- Effetti di relatedness semantica massimi
- LAN: Elaborazione morfosinttatica

500-900 ms: Integrazione sintattica e rianalisi
- P600: Integrazione struttura sintattica
- Rianalisi se necessario (garden-path)
- Risoluzione violazioni sintattiche
- Assegnazione ruoli tematici finali

> 900 ms: Processi post-interpretativi

- Integrazione discorsiva
- Aggiornamento modelli situazionali
- Computazione implicature

Formalmente, possiamo modellare il processamento come una cascata temporale di operazioni:

```
Process(Word_t) = Pipeline(
  Acoustic_Analysis(0-100ms),
  Lexical_Access(100-300ms),
  Semantic_Integration(300-500ms),
  Syntactic_Integration(500-900ms),
  Discourse_Update(>900ms)
)
```

dove ciascuna fase opera su rappresentazioni prodotte dalla fase precedente:

$$R_{t+1} = f_t(R_t, Input_t, Context_t)$$

con:
- R_t = rappresentazione allo stadio temporale t
- f_t = funzione computazionale dello stadio t
- $Input_t$ = input esterno allo stadio t
- $Context_t$ = contesto accumulato fino a t

Oscillazioni Neurali e Binding Temporale nell'Elaborazione Linguistica

Oltre alle componenti ERP, l'attività oscillatoria cerebrale in diverse bande di frequenza gioca ruoli specifici nel processamento linguistico. Le oscillazioni riflettono sincronizzazione ritmica di popolazioni neuronali e possono implementare meccanismi di binding e comunicazione inter-regionale.

Bande di frequenza rilevanti:

Delta (1-4 Hz): Tracking di flusso discorsivo

Funzione: Segmentazione di unità linguistiche grandi (frasi, enunciati)
Evidenza: Oscillazioni delta si agganciano (entrain) al ritmo prosodico del discorso
Picchi delta corrispondono a confini frasali

Meccanismo: Delta tracking permette predizione temporale di quando informazione critica arriverà

Formalizzazione:

$$Fase\_Delta(t) \propto Struttura\_Prosodica(Discorso(t))$$

Quando Fase_Delta = picco → Aspettativa di confine frasale

→ Preparazione per integrazione

Theta (4-8 Hz): Chunking di informazione, working memory

Funzione: Organizzazione temporale di chunks lessicali e sintattici
Mantenimento in working memory fonologica

Evidenza: Potenza theta aumenta durante:
- Mantenimento di liste di parole in memoria
- Processamento di frasi sintatticamente complesse
- Parsing di strutture gerarchiche

Localizzazione: Corteccia prefrontale, corteccia temporale

Meccanismo proposto:
Ogni ciclo theta (~150-200ms) corrisponde a un chunk di informazione
4-7 items possono essere mantenuti in fase theta (limite working memory)

Formalizzazione:
Capacità_WM = Frequenza_Theta · Durata_Mantenimento / Durata_Ciclo_Theta
≈ 5 Hz · 2s / 0,2s ≈ 7±2 items (Miller's magical number)

Alpha (8-13 Hz): Inibizione attiva, attenzione selettiva

Funzione: Soppressione di informazione irrilevante
Allocazione attenzionale a input linguistico vs altri canali

Evidenza: Potenza alpha aumenta in regioni che devono essere inibite
Alpha desincronizzazione (riduzione potenza) in regioni attivamente processanti

Esempio: Durante ascolto linguistico:
Corteccia uditiva: Alpha ↓ (desincronizzazione - attivo)
Corteccia visiva: Alpha ↑ (sincronizzazione - inibito)

Meccanismo: Alpha ritmica implementa gating dell'informazione
Fase alpha determina eccitabilità neuronale

Formalizzazione:
Eccitabilità(Regione, t) ∝ cos(Fase_Alpha(Regione, t) - π)
Massima eccitabilità quando fase = π (picco dell'onda)

Beta (13-30 Hz): Mantenimento di stato, feedback top-down

Funzione: Comunicazione feedback da corteccia frontale a regioni posteriori
Implementazione aspettative top-down

Evidenza: Accoppiamento beta tra regioni frontali e temporali durante predizione linguistica
Beta aumenta prima di parole altamente predicibili

Meccanismo: Beta implementa segnali di predizione che modulano processamento bottom-up

Formalizzazione:
Predizione_Top-Down(t) = β_Frontale(t) → Modulazione_Temporale(t+Δt)

Quando parola predetta arriva:
Mismatch piccolo → Beta mantiene
Mismatch grande → Beta reset → Oscillazioni gamma per risoluzione

Gamma (30-100 Hz): Binding locale, rappresentazioni attive

Funzione: Binding di features in rappresentazioni coerenti
Comunicazione feedforward rapida
Rappresentazione attiva di informazione semantica/fonologica

Evidenza:
- Burst gamma durante accesso lessicale
- Sincronizzazione gamma tra regioni che devono comunicare rapidamente
- Gamma aumenta per violazioni semantiche (N400) - riflette risoluzione conflitto

Localizzazione: Diffusa, particolarmente prominente in corteccia temporale durante linguaggio

Meccanismo:
Neuroni codificanti features correlate scaricano sincronicamente in gamma
Questa sincronizzazione binda features in oggetto/concetto unificato

Formalizzazione:
Rappresentazione(Concetto_C) = {Neurone_i | Fase_Gamma(Neurone_i) ≈ Fase_Gamma(Neurone_j) ∀i,j codificanti features di C}

Binding problem risolto via temporal synchrony:
Concetto_1: Neuroni $\{n_1, n_2, n_3\}$ sincronizzati in gamma fase φ_1
Concetto_2: Neuroni $\{n_4, n_5, n_6\}$ sincronizzati in gamma fase φ_2
Dove $\varphi_1 \neq \varphi_2$ (asynchronous gamma) → concetti distinti mantenuti separati

Cross-Frequency Coupling: Interazioni tra bande di frequenza

Mechanismo potente: Fase di oscillazioni lente modula ampiezza di oscillazioni rapide

Theta-Gamma Coupling durante sintassi:
Fase theta organizza when chunks sintattici sono processati

Ampiezza gamma codifica what (contenuto del chunk) quando fase theta è ottimale

Formalizzazione:
Ampiezza_Gamma(t) ∝ f(Fase_Theta(t))

Quando Fase_Theta(t) = φ_ottimale → Gamma_Burst
→ Processamento chunk sintattico

Delta-Alpha Coupling durante discorso:
Fase delta traccia struttura prosodica
Alpha modulato da delta implementa finestre attenzionali predittive

Evidenza empirica:
Misura: Phase-Amplitude Coupling (PAC)
PAC_Theta-Gamma aumenta durante processamento frasi complesse
PAC predice performance comportamentale in comprensione

Plasticità Neurale e Riorganizzazione dopo Lesione Cerebrale

Il sistema linguistico cerebrale dimostra notevole **plasticità**, ovvero capacità di riorganizzazione funzionale in risposta a lesione o esperienza. Questa plasticità è particolarmente evidente in casi di recupero da afasia e in popolazioni speciali come bambini con lesioni precoci.

Plasticità ontogenetica (durante sviluppo):

Periodo critico per linguaggio:

Ipotesi periodo critico (Lenneberg, 1967):
Esiste una finestra temporale ottimale per acquisizione linguistica
Approssimativamente: nascita - pubertà

Evidenza neurobiologica:
- Lateralizzazione emisferica per linguaggio si stabilizza durante infanzia
- Bambini con lesione emisferica sinistra precoce possono sviluppare linguaggio normale con emisfero destro
- Adulti con lesione equivalente mostrano afasia permanente

Formalizzazione:
Plasticità(età) = P_max · e^(-λ·età)

dove P_max = plasticità massimale alla nascita
λ = costante di decay (plasticità diminuisce esponenzialmente)

Età < 5 anni: Plasticità molto alta → Recupero completo possibile anche con lesioni estese
Età 5-12 anni: Plasticità moderata → Recupero parziale
Età > 12 anni: Plasticità ridotta → Recupero limitato, deficit residui tipici

Meccanismi di recupero in afasia adulta:

Tre pattern principali:

1. Riattivazione di regioni perilesionali:

Meccanismo: Tessuto cerebrale residuo nella vicinanza della lesione assume funzioni della regione danneggiata

Evidenza fMRI longitudinale:
Post-lesione acuta: Ippoattivazione perilesionale (diaschisi)
3-6 mesi post: Incremento progressivo attivazione perilesionale
Correlazione: Aumento attivazione perilesionale ↔ Miglioramento linguistico

Ottimale quando: Lesione non estensiva, margini di tessuto sano disponibili

Formalizzazione:
Funzione_Linguistica(t) ∝ Attivazione_Perilesionale(t)

Recupero = ∫[0→∞] Riorganizzazione_Perilesionale(t) dt

2. Reclutamento omologo emisferico:

Meccanismo: Emisfero destro (intatto) assume funzioni tipicamente lateralizzate a sinistra

Evidenza:
Pazienti con lesione LIFG estesa mostrano attivazione RIFG (omologo destro) durante compiti linguistici
Stimolazione magnetica transcranica (TMS) su RIFG → interferenza con produzione linguistica in questi pazienti

Controversia:
Alcuni studi: Attivazione emisfero destro è compensatoria e benefica
Altri studi: È maladattiva e interferisce con recupero ottimale

Ipotesi attuale (Heiss & Thiel):
Fase acuta: Reclutamento destro è compensatorio necessario
Fase cronica: Persistenza attivazione destra può essere maladattativa
Recupero ottimale richiede ri-lateralizzazione a sinistra

Formalizzazione:
Compensazione_Destra(t) benefica quando t < t_critico
Compensazione_Destra(t) maladattativa quando t > t_critico

dove t_critico ≈ 3-6 mesi post-lesione

3. Riorganizzazione di reti distribuite:

Meccanismo: La funzione danneggiata viene redistribuita attraverso una rete più estesa di regioni risparmiate

Evidenza da connettività funzionale:
Pre-lesione: Sintassi implementata da rete compatta (LIFG, pSTG, parietal)
Post-lesione (con lesione LIFG): Rete espansa include regioni frontali dorsali, temporale anteriore, corteccia prefrontale dorsolaterale

Interpretazione: La computazione sintattica viene "decomposta" e distribuita
Regioni addizionali contribuiscono sub-componenti del processo

Formalizzazione:
Pre-lesione: $Funzione\_F = \Sigma[i=1\rightarrow3]\ w\_i \cdot f\_i(Regione\_i)$ (rete compatta)

Post-lesione: $Funzione\_F = \Sigma[i=1\rightarrow8]\ w'\_i \cdot f\_i(Regione\_i)$ (rete espansa)

dove $\Sigma w'\_i \approx \Sigma w\_i$ (stessa funzione totale)
ma più regioni contribuiscono con pesi individuali minori

Fattori predittivi di recupero:

Predittori positivi (associati a miglior recupero):
- Età più giovane
- Lesione più piccola
- Risparmio di sostanza bianca (fibre di connessione)
- Attivazione perilesionale sinistra (non destra) in fase cronica
- Intensità terapia riabilitativa precoce
- Riserva cognitiva pre-morbosa alta (educazione, bilinguismo)

Modello predittivo empirico (semplificato):
$$Recupero = \beta_0 + \beta_1 \cdot (1/Età) + \beta_2 \cdot \log(1/Volume\_Lesione) + \beta_3 \cdot Integrità\_FA + \beta_4 \cdot Intensità\_Terapia + \beta_5 \cdot Riserva\_Cognitiva + \varepsilon$$

dove FA = fractional anisotropy (misura integrità sostanza bianca)

L'Acquisizione del Linguaggio nel Bambino: Meccanismi, Stadi e Teorie

Il Corso Temporale dell'Acquisizione: Tappe Universali dello Sviluppo Linguistico

L'acquisizione del linguaggio nel bambino segue una sequenza temporale notevolmente consistente attraverso culture e lingue diverse, suggerendo che il processo è guidato da

fattori maturazionali biologici oltre che da input ambientale. Possiamo delineare le tappe principali dello sviluppo linguistico dalla nascita alla prima infanzia.

Periodo Prelinguistico (0-12 mesi):

0-2 mesi: Fase riflessiva
Produzioni vocali: Pianto, suoni vegetativi (ruttini, starnuti)
Percezione: Discriminazione di contrasti fonetici universali
Preferenza per voce materna
Sensibilità a ritmo prosodico della lingua nativa

Evidenza sperimentale (High Amplitude Sucking Paradigm):
Neonati (giorni di vita) discriminano:
/ba/ vs /pa/ (contrasto voicing)
/ba/ vs /da/ (contrasto place of articulation)
Anche contrasti non-fonemici nella lingua nativa

Formalizzazione capacità percettive neonatali:
Discriminazione($Suono_1$, $Suono_2$) quando Distanza_Acustica(S_1, S_2) > θ_neonatale
dove θ_neonatale è basso (alta sensibilità)

2-4 mesi: Fase di cooing
Produzioni: "coo", "goo", suoni vocalici
Caratteristiche: Suoni prodotti nella parte posteriore del tratto vocale
Intonazione piacevole
Tipicamente in contesti di benessere

Funzione: Esplorazione capacità articolatorie
Controllo motorio respiratorio
Interazione sociale proto-comunicativa

6-10 mesi: Fase di babbling (lallazione)
Produzioni: Sequenze sillabiche ripetute: "bababa", "dadada", "mamama"

Canonical babbling (7-8 mesi):
Struttura: CV (consonante-vocale) ripetuta
Caratteristica acustica: Timing adulto-simile (transizioni rapide)

Variegated babbling (9-10 mesi):
Struttura: Sequenze sillabiche variate: "baduga", "manada"
Intonazione: Contorni prosodici complessi (jargon babbling)
Somiglianza: Inizia a rispecchiare patterns fonotattici della lingua nativa

Evidenza cross-linguistica:
Bambini di ambienti linguistici diversi producono babbling universalmente simile inizialmente
Ma verso 10 mesi, babbling diverge riflettendo struttura fonetica della lingua input

Esempio:
Bambini francofoni: Più sillabe con vocali anteriori /i, e/
Bambini arabofoni: Più consonanti faringali /ħ, ʕ/
Bambini giapponesi: Evitano clusters consonantici (giapponese ha struttura (C)V)

Formalizzazione convergenza verso lingua nativa:
Distribuzione_Fonemi_Babbling(età) → Distribuzione_Fonemi_Lingua_Nativa
quando età → 12 mesi

Convergenza = KL-Divergenza(Distr_Babbling || Distr_Lingua_Nativa)
KL → 0 quando età → 12 mesi

10-12 mesi: Comprensione di parole inizia
Prime comprensioni: Nomi propri ("Mamma", "Papà", nome bambino)
Routine sociali ("ciao", "no")
Nomi di oggetti familiari ("pappa", "acqua")

Vocabolario recettivo: 50-100 parole a 12 mesi

Gesti comunicativi emergono:
Deittici: Pointing (indicare), showing (mostrare)
Rappresentazionali: Simulare bere da tazza, fingere dormire

Evidenza sperimentale (Preferential Looking Paradigm):
Stimolo: "Guarda il cane!"
Setup: Due schermi mostrano cane vs gatto
Misura: Direzione e durata dello sguardo
Risultato: Bambini 10-12 mesi guardano significativamente più verso immagine nominata

Interpretazione: Associazione parola-referente già stabilita

Stadio Una-Parola / Olofrastico (12-18 mesi):

Produzioni: Singole parole isolate
Vocabolario produttivo:
12 mesi: 1-3 parole
15 mesi: 10-20 parole
18 mesi: 50-100 parole (inizio "vocabulary spurt")

Caratteristiche semantiche:
Tipi semantici tipici:
- Nomi di persone familiari (60%)
- Nomi comuni per oggetti (20%)
- Verbi di azione (10%)
- Aggettivi (5%)
- Parole funzionali (5%)

Concetti preferiti (Gentner's Natural Partitions Hypothesis):
Oggetti con confini chiari e forma distinta: "palla", "scarpa", "cucchiaio"
Non: masse omogenee o parti di oggetti

Overextension (sovraestensione):
Uso di parola per referenti oltre categoria convenzionale
Esempio: "cane" usato per tutti gli animali a quattro zampe
"palla" usato per tutti gli oggetti rotondi

Formalizzazione:
Categoria_Bambino(parola) ⊃ Categoria_Adulto(parola)
Criterio: Similarità percettiva saliente (forma, dimensione, movimento)

Underextension (sottoestensione):
Uso ristretto a esemplari specifici
Esempio: "cane" usato solo per il cane di famiglia, non altri cani

Formalizzazione:
Categoria_Bambino(parola) ⊂ Categoria_Adulto(parola)

Funzione pragmatica:
Enunciati olofrastici esprimono intenzioni comunicative complete

"Pappa!" può significare:
- Richiesta: "Voglio mangiare"
- Dichiarazione: "Questa è la pappa"
- Domanda: "È ora di mangiare?"

Disambiguazione tramite:
- Intonazione (ascendente per domanda, discendente per dichiarazione)
- Gesto accompagnatorio
- Contesto situazionale

Stadio Due-Parole (18-24 mesi):

Produzioni: Combinazioni di due parole
Vocabolario: 200-400 parole a 24 mesi

Esempi tipici:
"Mamma pappa" (agente-azione o possessore-posseduto)
"Più latte" (ricorrenza-oggetto)
"No nanna" (negazione-azione)
"Cane grande" (entità-attributo)
"Auto via" (entità-locazione)

Struttura semantica:
Relazioni semantiche espresse (Brown, 1973):

- Agente-Azione: "Papà mangia"
- Azione-Oggetto: "Mangia biscotto"
- Azione-Locazione: "Va fuori"
- Entità-Locazione: "Libro tavolo"
- Possessore-Posseduto: "Mio gioco"
- Entità-Attributo: "Palla rossa"
- Dimostrazione: "Quello cane"
- Ricorrenza: "Altro biscotto"
- Negazione: "No sonno"

Grammatica telegrafica:
Caratteristica: Omissione sistematica di elementi funzionali
Presenti: Parole di contenuto (nomi, verbi, aggettivi)
Assenti: Articoli, preposizioni, flessioni, copule, ausiliari

Adulto: "Il gatto è sul tavolo"
Bambino: "Gatto tavolo"

Adulto: "Voglio ancora del latte"
Bambino: "Più latte" o "Ancora latte"

Formalizzazione:
Enunciato_Bambino = Proietta({Parole_Contenuto} ∩ Enunciato_Target)

dove Proietta conserva ordine relativo ma elimina elementi funzionali

Ordine delle parole:
Non casuale! Riflette ordine canonico della lingua target

Italiano (SVO):
"Mamma mangia" (corretto)
*"Mangia mamma" (raro se inteso come dichiarazione)

Ma flessibilità maggiore che in adulto (topicalizzazione più libera)

Stadio Telegrafico Multi-Parola (24-30 mesi):

Lunghezza media enunciato (MLU - Mean Length of Utterance):
Calcolato in morfemi (non parole)

MLU(età_mesi) ≈ (età_mesi / 3) - 2

24 mesi: MLU ≈ 2,0-2,5 morfemi
30 mesi: MLU ≈ 2,5-3,5 morfemi
36 mesi: MLU ≈ 3,5-4,0 morfemi

Complessità sintattica aumenta:

"Papà legge libro"
"Voglio ancora biscotto"
"Bambina non piange"
"Dove andato cane?"

Emergenza morfologia:
Prime flessioni verbali:
Italiano: "-a" terza persona singolare: "Mangia", "Dorme"
Inglese: "-ing" progressivo: "eating", "running"

Plurali iniziano:
Regolari prima: "gatti", "cani"
Irregolari acquisiti per memorizzazione: "uomini" (non *"uomi")

Errori di sovraregolarizzazione (overgeneralization):
Applicazione di regola regolare a forme irregolari
Inglese: "goed" (invece di "went"), "mouses" (invece di "mice")
Italiano: "facete" (invece di "fate"), "andò" (invece di "andarono")

Interpretazione: Evidenza di acquisizione di regola produttiva
Prima fase: Forme memorizzate correttamente (irregolari)
Seconda fase: Sovraregolarizzazione (regola sovrapposta)
Terza fase: Sistema maturo (regola con eccezioni)

Curva a U:
Stadio I: "went" (corretto, memorizzato)
Stadio II: "goed" (scorretto, regola applicata)
Stadio III: "went" (corretto, sistema maturo)

Meccanismi di Apprendimento Lessicale: Il Problema dell'Induzione e Vincoli Cognitivi

L'acquisizione lessicale pone un problema logico profondo: come fanno i bambini a identificare i referenti corretti delle parole dato l'input radicalmente sottodeterminato? Questo è il **problema di Quine della traduzione radicale** applicato all'acquisizione.

Scenario:

Situazione: Adulto e bambino osservano un coniglio che salta
Adulto: "Guarda il coniglio!"
Bambino: Deve inferire cosa "coniglio" denota

Ipotesi logicamente possibili (infinite):
H_1: "Coniglio" denota l'animale intero
H_2: "Coniglio" denota le parti del coniglio
H_3: "Coniglio" denota il colore del coniglio
H_4: "Coniglio" denota il movimento di saltare

H_5: "Coniglio" denota la situazione coniglio-che-salta
H_6: "Coniglio" denota l'insieme di stadi temporali del coniglio
H_7: "Coniglio" denota la proprietà essere-un-coniglio
...
H_∞: Infinitamente molte altre interpretazioni logicamente coerenti

Problema dell'Induzione Lessicale: Dato un numero finito di esposizioni a una parola in contesti, come il bambino converge sull'interpretazione corretta dato l'insieme infinito di ipotesi consistenti con i dati?

Formalmente:

Input: {⟨Parola_w, Situazione_s_1⟩, ⟨w, s_2⟩, ..., ⟨w, s□⟩}

Ipotesi consistenti:
H_consistenti = {H | ∀i ∈ {1..n} : H è vera in s_i}

Tipicamente: |H_consistenti| = ∞

Problema: Come selezionare H_corretta ∈ H_consistenti?

Soluzione proposta: Vincoli di apprendimento (Learning Constraints)

I bambini non considerano tutte le ipotesi logicamente possibili ma applicano **vincoli innati** che restringono drasticamente lo spazio delle ipotesi candidate.

Vincolo 1: Whole Object Assumption (Assunzione Oggetto-Intero)

Formulazione (Markman, 1989):
Quando si impara un nuovo nome per un oggetto, assume che il nome denota l'oggetto intero, non parti o proprietà

Formalizzazione:
Data: Nuova parola w enunciata mentre bambino attende a oggetto O

Assunzione: Referente(w) = O (oggetto intero)
Non: Referente(w) = Parte(O) o Proprietà(O)

Predizioni empiriche:
✓ Bambini mappano rapidamente nuovi nomi a oggetti interi
✓ Apprendimento nomi di parti richiede evidenza disambiguante aggiuntiva (gesto indicante parte, contesto contrastivo)

Evidenza sperimentale:
Condizione: Bambino vede oggetto nuovo, sente "Questo è un dax"
Test: "Dammi il dax"
Risultato: Bambino seleziona oggetto intero (non parte come manico, non colore)

N = 200 bambini 2-3 anni, Accuracy = 87%

Vincolo 2: Taxonomic Assumption (Assunzione Tassonomica)

Formulazione:
Nomi di oggetti denotano categorie tassonomiche (oggetti dello stesso tipo)
Non: Raggruppamenti tematici o associazioni situazionali

Formalizzazione:
Estensione(Nome) = {x | Stesso_Tipo(x, Esemplare_Originale)}
Non: {x | Associato_Tematicamente(x, Esemplare_Originale)}

Predizione:
Dato: "Questo è un cane" (indicando un cane)

Estensione tassonomica: Altri cani, animali simili
Estensione tematica: Cibo per cani, guinzaglio, cuccia

Bambini preferiscono estensione tassonomica

Evidenza (Markman & Hutchinson, 1984):
Setup: Bambino impara nuova parola "dax" per oggetto (es. tazza speciale)
Test: Deve scegliere altro "dax" tra:
Opzione A: Oggetto tassonomicamente simile (altra tazza, forma diversa)
Opzione B: Oggetto tematicamente associato (piattino che va con tazza A)

Risultato:
Senza etichetta linguistica: 25% sceglie tassonomico
Con etichetta "dax": 75% sceglie tassonomico

Interpretazione: Etichette linguistiche attivano assunzione tassonomica

Vincolo 3: Mutual Exclusivity (Esclusività Mutua)

Formulazione:
Ogni oggetto ha un solo nome (a livello di base)
Se oggetto ha già un nome noto, nuova parola deve denotare qualcos'altro

Formalizzazione:
$\forall$ oggetto O, $\exists$!nome N : Livello_Base(N) $\wedge$ Denota(N, O)

Se Conosciuto($Nome_1$, $Oggetto_1$) $\wedge$ Nuova_Parola($Nome_2$):
$\rightarrow$ Referente($Nome_2$) $\neq$ $Oggetto_1$

Predizione (Fast Mapping):

Setup: Bambino conosce "tazza"
Tavolo ha tazza familiare + oggetto nuovo (pinza)
Istruzione: "Dammi il gombetto" (parola inventata)

Risultato: Bambino seleziona oggetto nuovo (pinza)

Ragionamento inferenziale del bambino:
(1) Tazza già ha nome "tazza"
(2) "Gombetto" è nuovo nome
(3) Per mutual exclusivity, "gombetto" ≠ "tazza"
(4) Quindi "gombetto" denota l'altro oggetto

Evidenza quantitativa:
Età 2 anni: 70% correct fast mapping
Età 3 anni: 90% correct fast mapping
Età 4 anni: 95% correct fast mapping

Meccanismo cognitivo:
Fast Mapping via Process of Elimination:
Referente(Parola_Nuova) = Unico_Referente_Senza_Etichetta_nel_Contesto

Vincolo 4: Shape Bias (Bias per Forma)

Formulazione (Landau, Smith & Jones):
Quando si generalizza nuovo nome per artefatti solidi, similarità di forma è dimensione critica
(Non colore, texture, dimensione)

Evidenza:
Training: "Questo è un dax" (mostrando oggetto a forma di T, blu, legno)

Test: Quale altro è un "dax"?
A: Forma T, rosso, metallo (forma uguale, altre proprietà diverse)
B: Forma L, blu, legno (forma diversa, altre proprietà uguali)
C: Forma T, dimensione molto diversa

Risultato: Bambini 2-4 anni scelgono A e C (basati su forma) > 85%

Formalizzazione:
Similarità_per_Categorizzazione_Bambini(x, y) ≈ Similarità_Forma(x, y)

Peso_Features:
w_forma = 0,8
w_colore = 0,1
w_texture = 0,05
w_dimensione = 0,05

Dominio-specificità:
Shape bias forte per: Artefatti solidi
Shape bias debole per: Sostanze, animali

"Questo sib" (sostanza gelatinosa verde):
Estensione basata su texture e sostanza, non forma

"Questo è un fep" (animale giocattolo):
Estensione basata su texture (tipo di animale) più che forma precisa

Apprendimento Statistico e Meccanismi di Tracciamento Distribuzionale

Oltre ai vincoli innati, i bambini possiedono potenti meccanismi di **apprendimento statistico** che estraggono pattern regolari dall'input linguistico attraverso tracciamento di distribuzioni e probabilità.

Segmentazione Lessicale via Probabilità Transizionali

Il problema della segmentazione:

Input al bambino: Flusso continuo di suoni senza pause chiare tra parole

Esempio flusso italiano:
/ilgattoèsultavolo/

Segmentazioni possibili:
Corretta: /il/ /gatto/ /è/ /sul/ /tavolo/
Scorretta: /ilgat/ /toe/ /sulta/ /volo/

Come il bambino identifica confini di parola senza supervisione esplicita?

Meccanismo: Tracciamento Probabilità Transizionali (TP)

Definizione:
TP(sillaba_B | sillaba_A) = P(B segue A) = Count(AB) / Count(A)

Principio:
TP alta all'interno di parole (sillabe co-occorrono frequentemente)
TP bassa ai confini di parole (sillabe più variabili)

Esempio:
Parole frequenti: "bambino", "cane", "mangia"

Frase: "il bambino mangia" (ripetuta frequentemente)

Calcolo TP:
TP(/bi/ | /bam/) = alto (sempre "bambino")

TP(/no/ | /bi/) = alto (sempre "bambino")
TP(/man/ | /no/) = variabile ("no mangia", "no corre", "no dorme")

Inferenza: Confine di parola dopo /no/ perché TP(/man/ | /no/) < TP_interno

Formalizzazione generale:
Confine_Parola_Prob(dopo sillaba_i) ∝ 1 / TP(sillaba_i+1 | sillaba_i)

Se TP < θ → confine probabile
dove θ è soglia appresa dinamicamente

Evidenza sperimentale seminale (Saffran, Aslin & Newport, 1996):

Paradigma:
Partecipanti: Neonati 8 mesi

Training:
Esposizione a 2 minuti di flusso sillabico continuo sintetizzato
Flusso contiene "parole" artificiali di 3 sillabe senza pause

"Parole": bidaku, padoti, golabu, tupiro (inventate)
Flusso: ...bidakupadotibidakugolabutupirobidakupadoti...

Proprietà statistiche:
TP all'interno di "parole" = 1,0 (deterministico)
TP tra "parole" = 0,33 (variabile, 3 parole possono seguire)

Test (Head-Turn Preference Procedure):
Tipo A: "Parole" sentite durante training (bidaku, padoti)
Tipo B: "Non-parole" mai sentite ma composte dalle stesse sillabe (dakupa, tigola)

Misura: Tempo di ascolto preferenziale

Risultato:
Neonati ascoltano più a lungo "parole" del training vs non-parole

Tempo_ascolto(Parole) = 7,2 sec > Tempo_ascolto(Non-parole) = 5,8 sec
p < 0,01

Interpretazione:
Dopo solo 2 minuti di esposizione, neonati hanno:
(1) Tracciato statistiche transizionali
(2) Segmentato flusso in unità ("parole")
(3) Riconosciuto quelle unità al test

Senza alcuna supervisione semantica o pause esplicite!

Generalizzazione: Apprendimento di Regolarità Fonologiche

Bambini usano distribuzione per apprendere fonotattiche (quali combinazioni fonemi sono permesse):

Italiano permette:
/st/ iniziale: "stella", "strada"
Non permette: /ts/ iniziale (*"tsella")

Bambino apprende:
P(/st/ | #inizio_parola) = alta → sequenza permessa
P(/ts/ | #inizio_parola) = 0 → sequenza proibita

Evidenza: Wug-test (Berko, 1958)
Bambini 4-5 anni generalizzano correttamente regole fonologiche a pseudo-parole nuove

"Questo è un /wʌg/. Ora ce ne sono due. Ci sono due ___?"
Risposta corretta: "/wʌgz/" (plurale regolare)

Performance: > 75% corretto

Interpretazione: Regola plurale /-z, -s, -ɪz/ appresa da distribuzione input

Parametri Sintattici e Triggering nel Framework Principles & Parameters

La teoria dei Principi e Parametri (Chomsky, 1981; 1986) propone che l'acquisizione sintattica consiste nel fissare valori di parametri universali basandosi sull'input.

Esempio Paradigmatico: Parametro Pro-Drop (Null Subject Parameter)

Definizione:
Le lingue variano se permettono omissione di pronomi soggetto in contesti finiti

Valori del parametro:
[+Pro-drop]: Soggetto pronominale può essere omesso
Lingue: Italiano, Spagnolo, Giapponese, Cinese, Arabo

[-Pro-drop]: Soggetto pronominale obbligatorio
Lingue: Inglese, Francese, Tedesco

Esempi:
Italiano [+Pro-drop]:
"Parlo italiano" (soggetto "io" omesso) ✓
"Io parlo italiano" (soggetto esplicito) ✓

Inglese [-Pro-drop]:
*"Speak English" (senza soggetto) ✗

"I speak English" (soggetto obbligatorio) ✓

Correlazioni parametriche (clustering):
[+Pro-drop] implica tipicamente:
- Morfologia verbale ricca (accordo persona/numero)
- Inversione soggetto-verbo libera
- That-trace effects assenti

[-Pro-drop] implica:
- Morfologia verbale povera
- Ordine soggetto-verbo rigido
- That-trace effects presenti

Triggering (innesco) del parametro:

Problema: Come il bambino determina il valore parametrico corretto dalla sua lingua?

Ipotesi Triggering:
Esistono "triggers" (inneschi) nell'input che disambiguano inequivocabilmente il valore

Per Pro-drop parameter:
Trigger positivo [+pro-drop]: Occorrenze di frasi finite senza soggetto esplicito
Input italiano: "Piove", "Mangio", "Dormi?"

Trigger negativo [-pro-drop]: Assenza di soggetti nulli + presenza sistematica di soggetti pronominali anche quando pragmaticamente ridondanti
Input inglese: "It rains", "I eat", "Do you sleep?"

Formalizzazione:
Parametro_Val = Funzione(Freq_Soggetti_Nulli_Input, Freq_Soggetti_Espliciti_Input)

Se Freq_Null > θ → [+Pro-drop]
Se Freq_Null ≈ 0 ∧ Freq_Explicit alta → [-Pro-drop]

Evidenza da acquisizione:

Predizione:
Bambini italiani fissano [+pro-drop] precocemente
Bambini inglesi fissano [-pro-drop] precocemente
Transizione è rapida (non graduale)

Evidenza empirica:
Studio longitudinale (Valian, 1991):

Bambini italiani (20-28 mesi):
Omissione soggetto: 70% frasi finite

Simile ad adulti italiani: 65%

Bambini inglesi (20-28 mesi):
Omissione soggetto: 10% frasi finite
(Errori residuali, correggono rapidamente)
Adulti inglesi: 0% (obbligatorio)

Età fissazione parametro: ~ 24-30 mesi per entrambe le lingue

Interpretazione:
Con input limitato, bambini convergono rapidamente su setting parametrico corretto
Supporta idea che parametri riducono drasticamente spazio ipotesi

Parametro Head-Direction (Direzionalità del Testa):

Definizione:
Posizione del testa (head) rispetto ai complementi nelle frasi

Valori:
[Head-Initial]: Testa precede complementi
Lingue: Inglese, Italiano, Francese (SVO, preposizioni)

[Head-Final]: Testa segue complementi
Lingue: Giapponese, Turco, Coreano (SOV, postposizioni)

Esempi:
VP (Verb Phrase):
Inglese [Head-initial]: [VP eat [NP an apple]]
Giapponese [Head-final]: [VP [NP ringo-o] taberu]
mela-ACC mangiare

PP (Prepositional/Postpositional Phrase):
Italiano: [PP in [NP casa]] (preposizione)
Giapponese: [PP [NP ie-ni]] (postposizione)
casa-a

Trigger:
Osservazione ordine relativo verbo-oggetto e preposizione-nome

Acquisizione:
Bambini fissano correttamente entro 24 mesi
Errori di ordine verbo-oggetto sono rari (< 5%)

Evidenza: Nessun bambino inglese produce consistentemente SOV
Nessun bambino giapponese produce consistentemente SVO

Periodo Critico e Sensibilità Maturazionale

L'**Ipotesi del Periodo Critico** (Critical Period Hypothesis, CPH) propone che esiste una finestra temporale limitata durante la quale l'acquisizione linguistica può avvenire completamente e senza sforzo, dopo la quale l'acquisizione diventa significativamente più difficile o impossibile.

Evidenza da privazione linguistica:

Caso Genie (Curtiss, 1977):
Storia: Bambina severamente abusata, isolata linguisticamente
Scoperta a età 13 anni con virtualmente zero esposizione linguistica

Intervento: Training intensivo linguistico per anni

Risultato:
Vocabolario: Espansione sostanziale (centinaia di parole apprese)
Sintassi: Rimasta severamente compromessa

Produzioni tipiche età 18:
"Applesauce buy store" (Telegrafico, agrammatico)
*"Father hit leg" (Per "Father hit me on the leg")

Morfologia: Quasi totalmente assente
Nessun uso consistente di:
- Flessioni verbali (-ed, -ing)
- Determinanti (the, a)
- Ausiliari (is, has, will)

Interpretazione:
Oltre periodo critico → sintassi non più acquisibile normalmente
Lessico meno vincolato temporalmente (acquisizione possibile tardivamente)

Evidenza da acquisizione L2 (seconda lingua):

Studio classico (Johnson & Newport, 1989):

Partecipanti: Immigrati cinesi e coreani in USA
Variabile indipendente: Età di arrivo in USA (Age of Arrival, AoA)
Range: 3-39 anni

Variabile dipendente: Competenza grammaticale inglese
Test: Giudizi grammaticalità di frasi inglesi (216 items)

Risultati:
Correlazione AoA-Performance: r = -0,77 (forte, negativa)

AoA < 7 anni: Performance = 269/276 (97%) - livello nativo
AoA 8-10: 256/276 (93%)
AoA 11-15: 235/276 (85%)
AoA 17-39: 210/276 (76%) - plateau, alta variabilità

Funzione AoA-Performance:
Performance(AoA) = {
98% - 2%·AoA se AoA ≤ 7
95% - 1,5%·AoA se 7 < AoA ≤ 15 (declino lineare)
75% ± 10% se AoA > 15 (plateau, no ulteriore declino)
}

Interpretazione:
Finestra ottimale: nascita-7 anni (acquisizione nativa possibile)
Periodo sensibile: 7-15 anni (declino graduale)
Post-critico: >15 anni (competenza limitata, variabile)

Basi neurobiologiche del periodo critico:

Maturazione cerebrale e plasticità:

Nascita-3 anni: Sinaptogenesi massiva
- Densità sinaptica: 150% livello adulto
- Plasticità: Massimale
- Lateralizzazione: Non ancora completa

3-7 anni: Pruning sinaptico inizia
- Potatura sinapsi non-utilizzate
- Rafforzamento sinapsi usate frequentemente
- Stabilizzazione circuiti neurali per L1

7-Pubertà: Pruning continua, lateralizzazione si completa
- Densità sinaptica diminuisce verso livello adulto
- Specializzazione emisferica sinistra per linguaggio
- Riduzione plasticità cross-modale

Post-pubertà: Struttura relativamente fissata
- Densità sinaptica = livello adulto (stabile)
- Plasticità molto ridotta
- Nuovi apprendimenti linguistici richiedono circuiti alternativi

Markers neurobiologici di chiusura periodo critico:
- Mielinizzazione completa tratti linguistici: ~12-15 anni
- Stabilizzazione recettori NMDA (plasticità): ~pubertà
- Livelli ormoni (testosterone, estrogeni): cambio a pubertà modula plasticità

Formalizzazione computazionale:

Modello di Elman (1993) - Reti Neurali e "Less is More":

Ipotesi: Limitazioni cognitive iniziali (working memory ridotta) aiutano acquisizione

Simulazione:
Rete neurale feed-forward addestrata su grammatica artificiale complessa

Condizione A: "Adulto" - tutta la complessità grammaticale da inizio
Risultato: Fallimento apprendimento (overwhelmed)

Condizione B: "Bambino" - complessità aumenta gradualmente
Inizialmente: Solo frasi brevi, semplici
Gradualmente: Frasi più lunghe, complesse introdotte
Risultato: Apprendimento successo

Interpretazione:
"Less is More" - limitazioni iniziali filtrano input, focalizzando apprendimento su pattern semplici fondamentali
Base solida di pattern semplici permette poi scaffold per complessità maggiore

Formalizzazione capacità di apprendimento:
$L(età) = L_{max} \cdot [1 - \exp(-\lambda \cdot Maturazione(età))] \cdot \exp(-\mu \cdot età)$

Termine 1: $[1 - \exp(-\lambda \cdot Mat)]$ - beneficio da maturazione cognitiva (aumenta inizialmente)
Termine 2: $\exp(-\mu \cdot età)$ - costo da riduzione plasticità (diminuisce con età)

Massimo a $età_{ottimale}$ dove $dL/d(età) = 0$
Tipicamente $età_{ottimale} \approx 3\text{-}7$ anni

Verso una Teoria Unificata

Proseguendo l'analisi sistematica della natura della comprensione linguistica che abbiamo sviluppato attraverso i capitoli precedenti, diventa ora imperativo affrontare la questione teorica fondamentale che sottende l'intero edificio concettuale: è possibile formulare una teoria unificata della competenza linguistica che integri coerentemente i risultati della semantica formale, della pragmatica cognitiva, della neurolinguistica, della psicolinguistica sperimentale e degli studi sull'acquisizione in un framework teorico omogeneo e matematicamente rigoroso? Questa domanda, che potremmo definire la questione dell'unificazione teorica in linguistica, rappresenta forse la sfida intellettuale più ambiziosa e al contempo più necessaria per il progresso della disciplina nel ventunesimo secolo.

La difficoltà intrinseca di questo progetto unificatorio emerge chiaramente quando consideriamo le tensioni teoriche profonde tra i diversi paradigmi che abbiamo esaminato. Da un lato, la tradizione della semantica formale, che affonda le sue radici nel programma logicista di Frege, Russell e Carnap e che trova la sua espressione contemporanea più

sofisticata nella semantica di Montague e nelle sue estensioni dinamiche e intensionali, concepisce il significato linguistico come una relazione matematicamente precisabile tra espressioni linguistiche e strutture del mondo, tipicamente modellate attraverso la teoria dei modelli e la logica di ordine superiore. In questo paradigma, la competenza semantica di un parlante consiste fondamentalmente nella conoscenza implicita o tacita di funzioni compositive che mappano espressioni sintatticamente strutturate su condizioni di verità o, più generalmente, su contenuti proposizionali valutabili rispetto a mondi possibili.

La nozione centrale è quella di composizionalità semantica, secondo cui il significato di un'espressione complessa è sistematicamente determinato dai significati delle sue parti costituenti e dal modo in cui queste parti sono sintatticamente combinate. Formalmente, se interpretiamo il significato di un'espressione α come una funzione ⟦α⟧ che mappa α alla sua denotazione in un modello M, allora il principio di composizionalità può essere enunciato come l'esistenza di funzioni compositive f tali che per ogni espressione complessa della forma [α β γ], dove β è un operatore sintattico che combina α e γ, abbiamo:

⟦[α β γ]⟧ = f_β(⟦α⟧, ⟦γ⟧)

Questa architettura formale ha il pregio della precisione matematica e della generalità, permettendoci di derivare rigorosamente le proprietà inferenziali degli enunciati dalle loro strutture semantiche composte.

Dall'altro lato, il paradigma cognitivo-funzionalista, esemplificato dalla Teoria della Pertinenza di Sperber e Wilson, dalle teorie costruzioniste di Goldberg e Tomasello, e dalle teorie della metafora concettuale di Lakoff e Johnson, enfatizza che il linguaggio è fondamentalmente uno strumento per la comunicazione umana, e che la comprensione linguistica non può essere ridotta alla computazione di condizioni di verità ma richiede processi inferenziali pragmatici complessi che integrano informazione linguistica con conoscenza enciclopedica, teoria della mente, e principi di razionalità comunicativa. In questa prospettiva, ciò che viene codificato convenzionalmente dalle espressioni linguistiche è radicalmente sottodeterminato rispetto a ciò che viene comunicato: il contenuto semantico letterale fornisce soltanto uno scheletro schematico che deve essere arricchito pragmaticamente attraverso processi inferenziali guidati da aspettative di pertinenza per produrre l'explicature, ovvero la proposizione completamente articolata che costituisce il contenuto esplicito comunicato.

La comprensione linguistica diventa quindi un processo ostensivo-inferenziale in cui l'ascoltatore inferisce le intenzioni comunicative del parlante basandosi non su regole compositive deterministiche ma su euristiche probabilistiche che massimizzano il rapporto tra effetti cognitivi positivi e sforzo di elaborazione. Formalmente, potremmo modellare questo processo come una funzione di ottimizzazione: dato un enunciato U in un contesto C, l'interpretazione ottimale I* è quella che massimizza la funzione obiettivo:

Pertinenza(I, C) = EffettiCognitiviPositivi(I, C) / SforzoDiElaborazione(I, C)

soggetto al vincolo che Pertinenza(I*, C) ≥ θ dove θ è una soglia di accettabilità determinata dalla presunzione di pertinenza ottimale comunicata ostensivamente dall'enunciato stesso.

Integrazione delle Prospettive Formale e Cognitiva

Queste due prospettive sembrano operare a livelli esplicativi differenti e con obiettivi teorici parzialmente divergenti. La semantica formale aspira a specificare le condizioni di verità degli enunciati in modo composizionale e context-independent, trattando il contesto soltanto come una funzione di saturazione di parametri indicali e di risoluzione di ambiguità. La pragmatica cognitiva, al contrario, enfatizza la centralità del contesto e dei processi inferenziali nella determinazione del contenuto comunicato, e tratta il contenuto semantico codificato come un mero punto di partenza per l'interpretazione piuttosto che come il significato completo.

Una possibile via di integrazione emerge quando riconosciamo che semantica formale e pragmatica cognitiva non sono necessariamente teorie concorrenti della stessa fenomenologia ma piuttosto teorie complementari che si concentrano su aspetti diversi dell'architettura cognitiva della comprensione linguistica. La semantica formale può essere concepita come una teoria della conoscenza linguistica implicita, ovvero della competenza semantica intesa come sistema di regole compositive internamente rappresentate che generano rappresentazioni semantiche schematiche a partire da strutture sintattiche. Queste rappresentazioni semantiche, che potremmo chiamare forme logiche arricchite o proto-proposizioni, costituiscono l'output del sistema compositivo semantico e l'input per i processi pragmatici di interpretazione. La pragmatica cognitiva, dal canto suo, è una teoria della performance interpretativa, ovvero dei meccanismi inferenziali online che operano su queste forme logiche per produrre interpretazioni contestualmente appropriate attraverso arricchimento, disambiguazione e derivazione di implicature.

Architettura Computazionale Multi-Stadio

Formalizziamo questa architettura integrata con maggiore precisione. Supponiamo che l'elaborazione di un enunciato linguistico U proceda attraverso una pipeline computazionale strutturata in stadi successivi, ciascuno caratterizzato da un livello rappresentazionale specifico e da operazioni computazionali distinte.

Stadio 1: Percezione e Riconoscimento Fonologico

Al livello più basso, abbiamo l'input acustico o ortografico grezzo I_acustico, che deve essere convertito in una rappresentazione fonologica o grafematica discreta attraverso processi di percezione del parlato o lettura visiva. Questo primo stadio può essere modellato come una funzione di riconoscimento percettivo Φ_percezione che mappa il segnale fisico continuo su una sequenza di unità fonologiche discrete:

$\Phi\_percezione: I\_acustico \rightarrow SequenzaFonemi = \langle f_1, f_2, ..., f\_n \rangle$

dove ciascun f_i è un fonema o un cluster di features fonologiche.

Già a questo livello, dobbiamo riconoscere che la mappatura non è deterministica ma probabilistica, dato che il segnale acustico è intrinsecamente ambiguo a causa di variabilità nella pronuncia, rumore ambientale, e sovrapposizione temporale di features acustiche. Una modellazione più accurata richiede quindi che Φ_percezione produca non una singola sequenza fonemica ma una distribuzione di probabilità su possibili sequenze:

$\Phi\_percezione: I\_acustico \rightarrow P(SequenzaFonemi \mid I\_acustico)$

Questa distribuzione può essere implementata attraverso modelli acustici probabilistici come Hidden Markov Models o, più recentemente, attraverso reti neurali profonde addestrate su vasti corpus di parlato annotato. La rete neurale apprende una funzione complessa che mappa il segnale acustico rappresentato come spettrogramma o altra rappresentazione tempo-frequenza su distribuzioni di probabilità su sequenze fonemiche:

$P(f_1, ..., f_n \mid I_{acustico}) = NeuralNetwork_\theta(Spectrogram(I_{acustico}))$

dove θ rappresenta i parametri della rete neurale ottimizzati durante il training.

Stadio 2: Accesso Lessicale e Segmentazione

Una volta ottenuta la rappresentazione fonologica, il sistema lessicale deve segmentare la sequenza fonemica in unità lessicali discrete e recuperare le rappresentazioni lessicali corrispondenti dalla memoria lessicale a lungo termine. Questo processo di accesso lessicale può essere concepito come una funzione Φ_lessico che mappa sequenze di fonemi su liste di items lessicali insieme alle loro proprietà sintattiche, semantiche e morfologiche:

$\Phi_{lessico}: SequenzaFonemi \rightarrow Lista\langle Lessema_1, Lessema_2, ..., Lessema_m\rangle$

dove ciascun Lessema_i è una struttura complessa contenente minimalmente:

- Categoria sintattica (N, V, A, P, ...)
- Forma fonologica
- Semantica lessicale (componente intensionale)
- Proprietà di subcategorizzazione
- Frame argomentale
- Proprietà morfologiche

Anche questo processo è intrinsecamente probabilistico e competitivo. Date le proprietà di attivazione parallela e interattiva del sistema lessicale umano, dimostrate estensivamente attraverso paradigmi sperimentali di priming lessicale e eye-tracking, l'accesso lessicale non consiste nella selezione seriale di un singolo lessema ma piuttosto nell'attivazione graduata di un insieme di candidati lessicali la cui attivazione è modulata da molteplici fattori.

Formalmente, possiamo modellare questo come una funzione che assegna pesi di attivazione:

$\Phi_{lessico}: SequenzaFonemi \rightarrow \{\langle Lessema_i, peso(Lessema_i)\rangle \mid i = 1..m\}$

dove $peso(Lessema_i) \in [0, 1]$ riflette il grado di attivazione del lessema i dato l'input fonologico e il contesto.

L'attivazione di ciascun lessema è determinata da molteplici fattori che contribuiscono in modo interattivo:

$peso(L) = f(SimilaritàFonologica(Input, L), FrequenzaLessicale(L), RecencyEffetti(L), ContestoSemantico(L), NeighborhoodDensity(L))$

dove:

- SimilaritàFonologica quantifica quanto l'input fonologico corrisponde alla forma fonologica del lessema
- FrequenzaLessicale è la frequenza logaritmica del lessema nel linguaggio
- RecencyEffetti cattura priming da usi recenti
- ContestoSemantico cattura vincoli semantici top-down
- NeighborhoodDensity è il numero di lessemi fonologicamente simili che competono

Stadio 3: Parsing Sintattico

Procedendo verso livelli rappresentazionali superiori, il parser sintattico deve costruire una struttura sintattica gerarchica che specifica le relazioni grammaticali tra gli elementi lessicali identificati. Nella tradizione generativista, questa struttura è tipicamente concepita come un albero sintattico conforme ai principi della grammatica universale, con nodi etichettati da categorie sintattiche e relazioni di dominanza e precedenza che codificano la struttura costituente.

Il processo di parsing può essere formalizzato come una funzione Φ_sintassi che mappa sequenze di items lessicali su alberi sintattici:

$\Phi\_sintassi: Lista\langle Lessema_1, ..., Lessema\_m\rangle \rightarrow AlberoSintattico\ T$

Un albero sintattico T può essere rappresentato formalmente come una tupla:

$T = \langle N, E, \lambda, \lessdot \rangle$

dove:

- N è l'insieme di nodi
- $E \subseteq N \times N$ è l'insieme di archi diretti (relazioni di dominanza)
- $\lambda: N \rightarrow$ Categorie è una funzione di etichettatura che assegna categorie sintattiche ai nodi
- $\lessdot$ è una relazione d'ordine totale su nodi sorelle (precedenza lineare)

Anche qui, il non-determinismo emerge perché le grammatiche del linguaggio naturale generano tipicamente ambiguità strutturali: una singola sequenza di parole può essere compatibile con molteplici analisi sintattiche alternative. Il parser umano deve quindi navigare questo spazio di analisi possibili, e la ricerca psicolinguistica ha rivelato che questo processo è guidato da euristiche e preferenze strutturali probabilistiche.

Una modellazione computazionale accurata richiede quindi l'adozione di grammatiche probabilistiche, come le Probabilistic Context-Free Grammars (PCFGs), in cui ogni regola di produzione G ha associata una probabilità P(G). La probabilità di un albero sintattico T è il prodotto delle probabilità delle regole usate nella sua derivazione:

$P(T) = \prod_{i=1}^{k} P(G_i)$

dove $\{G_1, ..., G\_k\}$ è l'insieme delle regole applicate per derivare T.

Il parser cerca l'albero più probabile data la sequenza di parole:

$T^* = \text{argmax}_T P(T \mid w_1, ..., w_n)$

Attraverso il teorema di Bayes:

$P(T \mid w_1, ..., w_n) \propto P(w_1, ..., w_n \mid T) \cdot P(T)$

Dove $P(w_1, ..., w_n \mid T) = 1$ se le foglie di T corrispondono esattamente a $w_1, ..., w_n$, e 0 altrimenti. Quindi:

$T^* = \text{argmax}_{T \text{ compatibile}} P(T)$

Elaborazione Incrementale e Predizione

Un aspetto cruciale del parsing umano che questa caratterizzazione cattura soltanto parzialmente è l'incrementalità: gli umani non aspettano di ricevere l'enunciato completo prima di iniziare il parsing, ma piuttosto costruiscono strutture sintattiche incrementalmente parola per parola man mano che l'input arriva. Inoltre, il sistema non si limita a reagire passivamente all'input ma genera attivamente predizioni circa le parole e strutture future basandosi sul contesto sintattico e semantico già elaborato.

Formalmente, possiamo modellare il parsing incrementale come un processo che mantiene una distribuzione di probabilità su strutture sintattiche parziali che viene aggiornata sequenzialmente man mano che ciascuna nuova parola arriva:

$P(T_{partial} \mid w_1, ..., w_i) \rightarrow P(T_{partial}' \mid w_1, ..., w_i, w_{i+1})$

dove $T_{partial}$ è un albero sintattico parziale che copre le prime i parole, e $T_{partial}'$ è un'estensione che incorpora la parola i+1.

L'aggiornamento può essere implementato attraverso algoritmi di parsing probabilistico incrementale come il beam search parser, che mantiene in ogni passo le k analisi parziali più probabili (il "beam"):

$Beam_i = \{T_1^{(i)}, ..., T_k^{(i)}\}$ tale che $P(T_j^{(i)} \mid w_1, ..., w_i)$ sono le k probabilità più alte

Quando la parola w_{i+1} arriva, ciascuna analisi parziale nel beam viene estesa in tutti i modi possibili conformi alla grammatica, e le k estensioni più probabili formano il nuovo beam:

$Beam_{i+1} = \text{top}_k(\{\text{Estensione}(T, w_{i+1}) \mid T \in Beam_i\})$

La generazione di predizioni può essere modellata attraverso la distribuzione di probabilità marginale su possibili parole future dato il prefisso elaborato fino a ora:

$P(w_{i+1} \mid w_1, ..., w_i) = \sum_T P(w_{i+1}, T \mid w_1, ..., w_i) = \sum_T P(w_{i+1} \mid T, w_1, ..., w_i) \cdot P(T \mid w_1, ..., w_i)$

Questa distribuzione predittiva quantifica le aspettative del parser circa quali parole sono probabili in quella posizione sintattica, e queste aspettative influenzano l'elaborazione della parola effettivamente ricevuta: parole altamente predette sono processate più rapidamente (ridotto N400 in ERP), mentre parole impredette causano costi di elaborazione aumentati e possono innescare rianalisi.

Stadio 4: Composizione Semantica

L'output del parser sintattico costituisce l'input per il modulo di interpretazione semantica compositiva, che applica regole di composizione semantica per derivare una rappresentazione del significato della frase. In un framework montagoviano standard, questo processo consiste nell'applicazione ricorsiva di una funzione di interpretazione $[\![\cdot]\!]$ che assegna denotazioni a ciascun nodo dell'albero sintattico basandosi sulle denotazioni dei suoi figli e sul modo sintattico di combinazione, implementando così il principio di composizionalità.

Per un nodo N con figli N_1 e N_2 combinati attraverso l'operazione sintattica β, abbiamo:

$[\![N]\!] = f\_\beta([\![N_1]\!], [\![N_2]\!])$

dove f_β è la funzione semantica associata all'operazione sintattica β.

Nella semantica di tipo, questa funzione è tipicamente l'applicazione funzionale quando uno dei costituenti è di tipo funzionale compatibile con il tipo dell'altro costituente. Per esempio, se N_1 ha tipo $\langle e,t\rangle$ (funzione da entità a valori di verità, tipicamente un predicato verbale) e N_2 ha tipo e (entità, tipicamente un argomento nominale), allora:

$[\![[N_1\ N_2]]\!] = [\![N_1]\!]([\![N_2]\!])$

Questa applicazione funzionale produce un valore di verità (tipo t) che rappresenta se la proprietà denotata da N_1 si applica all'entità denotata da N_2.

Il risultato finale di questo processo compositivo bottom-up è una rappresentazione semantica della frase completa, che nella tradizione vero-condizionale corrisponde a una proposizione o, più precisamente nella semantica intensionale, a una funzione da indici contestuali (mondi possibili, tempi, parlanti, etc.) a valori di verità:

$[\![\text{Frase}]\!]$: Contesti → {Vero, Falso}

o equivalentemente, a un insieme di contesti (l'insieme di contesti che rendono la frase vera):

$[\![\text{Frase}]\!] = \{c \in \text{Contesti} \mid \text{Frase è vera in } c\}$

Sottodeterminazione Semantica e Arricchimento Pragmatico

È importante sottolineare però, come abbiamo enfatizzato ripetutamente nei capitoli precedenti, questa rappresentazione semantica compositiva è radicalmente

sottodeterminata rispetto al contenuto comunicato effettivamente. La forma logica derivata compositivamente contiene molteplici tipi di sottodeterminazione che richiedono risoluzione pragmatica:

Tipi di Sottodeterminazione Semantica

1. Sottodeterminazione Indicale

La forma logica contiene parametri indicali che devono essere saturati con valori contestuali specifici. Gli indicali puri come "io", "tu", "qui", "ora" hanno significati caratteriali costanti (character nel senso di Kaplan) ma riferimenti variabili dipendenti dal contesto:

⟦"io"⟧^{c,w} = parlante(c) ⟦"tu"⟧^{c,w} = destinatario(c)
⟦"qui"⟧^{c,w} = luogo(c) ⟦"ora"⟧^{c,w} = tempo(c)

dove c è un contesto che specifica almeno la tupla ⟨parlante, destinatario, tempo, luogo⟩ e w è un mondo possibile.

2. Sottodeterminazione Referenziale

Espressioni referenziali come pronomi anaforici, dimostrativi, e descrizioni definite richiedono risoluzione referenziale contestuale. Un pronome come "lei" codifica soltanto features grammaticali minime (femminile, singolare, terza persona) ma il suo referente deve essere recuperato dal contesto discorsivo:

⟦"lei"⟧^{c,w} = ι x [Femminile(x) ∧ Singolare(x) ∧ Saliente_in(x, c)]

dove Saliente_in(x, c) è una funzione contestualmente determinata che identifica il referente più saliente che soddisfa i vincoli grammaticali.

3. Sottodeterminazione di Restrizioni di Dominio

Quantificatori e altri operatori hanno restrizioni di dominio che sono raramente specificate esplicitamente nella forma superficiale:

"Tutti hanno superato l'esame"

La forma logica contiene un quantificatore universale con dominio non ristretto:

∀x [Persona(x) → Superato_esame(x)]

Ma pragmaticamente questo è interpretato con dominio ristretto:

∀x [Persona(x) ∧ Studente_del_corso(x) → Superato_esame(x)]

La restrizione contestuale "Studente_del_corso" è arricchimento pragmatico libero non codificato semanticamente.

4. Sottodeterminazione Temporale

Relazioni temporali espresse da tempi verbali sono spesso schematiche e richiedono specificazione pragmatica. Il passato prossimo italiano "ho fatto colazione" codifica soltanto:

$\exists t$ [Fare_colazione(io, t) $\wedge$ t < ora]

Ma è pragmaticamente arricchito a:

$\exists t$ [Fare_colazione(io, t) $\wedge$ t < ora $\wedge$ t $\in$ Mattina_di(Oggi)]

5. Sottodeterminazione di Predicati e Relazioni

Molti predicati sono semanticamente vaghi o generici e richiedono precisazione contestuale. "Grande" è relativo a una classe di confronto:

"Questo è un topo grande"

Non significa semplicemente Grande(topo) ma:

Grande_per_uno(topo, Classe_Confronto_Topi)

La classe di confronto deve essere determinata pragmaticamente.

L'Interfaccia Semantica-Pragmatica

Il gap tra forma logica e proposizione completamente articolata è il dominio della pragmatica. Possiamo modellare l'interfaccia semantica-pragmatica introducendo una funzione di arricchimento pragmatico Φ_pragmatica che prende come input la forma logica sottodeterminata LF insieme al contesto conversazionale C e produce come output una proposizione completamente specificata che costituisce l'explicature:

Φ_pragmatica: ⟨LF, C⟩ → Proposizione

Questa funzione implementa processi multipli di arricchimento che operano in parallelo o sequenzialmente:

Processo 1: Saturazione di Indicali

Saturazione: ⟨LF, C⟩ → LF'

Sostituisce parametri indicali con valori contestuali:

- io* → parlante(C)
- tu* → destinatario(C)
- ora* → tempo(C)
- qui* → luogo(C)

Processo 2: Risoluzione Referenziale

RisoluzioneRef: ⟨LF', C⟩ → LF''

Per ogni espressione referenziale sottodeterminata α in LF':

- Identificare insieme di referenti candidati: Candidati(α) = {x | x soddisfa features di α}
- Selezionare referente più saliente: ref(α) = argmax_{x∈Candidati} Salienza(x, C)
- Sostituire α con ref(α) in LF'

La funzione di salienza Salienza(x, C) è una combinazione pesata di molteplici fattori:

$Salienza(x, C) = w_1 \cdot Recency(x, C) + w_2 \cdot Topicalità(x, C) + w_3 \cdot Prominence\_Grammaticale(x, C) + w_4 \cdot Compatibilità\_Semantica(x, C)$

dove:

- Recency: quanto recentemente x è stato menzionato
- Topicalità: quanto x è topic corrente del discorso
- Prominence_Grammaticale: ruolo sintattico di menzione precedente (soggetto > oggetto)
- Compatibilità_Semantica: quanto x è semanticamente compatibile con predicato corrente

Processo 3: Arricchimento Libero

ArricchimentoLibero: ⟨LF", C⟩ → LF"'

Aggiunge materiale concettuale non codificato semanticamente ma richiesto per pertinenza:

- Restrizioni temporali: "ieri" → "ieri sera" (specificazione)
- Restrizioni di dominio: "tutti" → "tutti gli studenti del corso"
- Precisazione predicati: "grande" → "grande per un topo"
- Inferenze bridging: "Giovanni entrò nella stanza. Il soffitto era alto" → implicatura bridging che la stanza ha un soffitto

Processo 4: Assegnamento di Scope

ScopeAssignment: ⟨LF"', C⟩ → LF""

Risolve ambiguità di scope tra quantificatori, negazioni, e operatori modali basandosi su preferenze di scope e compatibilità semantica-pragmatica:

"Ogni studente non ha superato qualche esame"

Readings possibili:

- $\forall x \exists y$ ¬Superato(x,y): Per ogni studente, c'è qualche esame che non ha superato (wide universal)
- $\neg \forall x \exists y$ Superato(x,y): Non è vero che ogni studente ha superato qualche esame (wide negation)
- $\exists y \forall x$ ¬Superato(x,y): C'è un esame che nessun studente ha superato (wide existential)

La selezione è guidata da:

- Preferenza scope superficiale (left-to-right)
- Compatibilità plausibilità mondiale
- Massimizzazione pertinenza

Teoria della Pertinenza come Principio Guida

La teoria della pertinenza fornisce una caratterizzazione precisa di come questa funzione di arricchimento pragmatico opera. Secondo il Principio Comunicativo di Pertinenza, ogni atto di comunicazione ostensiva comunica una presunzione della propria pertinenza ottimale, guidando l'ascoltatore a seguire un percorso di minimo sforzo nell'elaborazione interpretativa, arrestandosi appena un'interpretazione che soddisfa le aspettative di pertinenza viene identificata.

Formalmente, l'ascoltatore esplora lo spazio delle interpretazioni possibili $I \in$ Spazio_Interpretazioni(LF, C) in ordine di accessibilità crescente, dove l'accessibilità di un'interpretazione è inversamente correlata allo sforzo cognitivo richiesto per derivarla, e seleziona la prima interpretazione I^* tale che:

$$\text{Pertinenza}(I^*, C) \geq \theta$$

dove θ è la soglia di pertinenza attesa.

La pertinenza di un'interpretazione I in un contesto C è definita come il rapporto tra effetti cognitivi positivi e sforzo di elaborazione:

$$\text{Pertinenza}(I, C) = \text{EffettiCognitivi}(I, C) / \text{Sforzo}(I, C)$$

Effetti Cognitivi Positivi

Gli effetti cognitivi includono tre tipi principali:

1. **Implicazioni Contestuali**: Nuove conclusioni derivabili dalla combinazione di I con assunzioni già presenti in C:

   $$\text{ImplicazioniContestuali}(I, C) = \{A \mid (C \cup \{I\}) \vdash A \wedge C \nvdash A \wedge \{I\} \nvdash A\}$$

2. **Rafforzamento**: Incremento della forza epistemica di assunzioni esistenti:

   $$\text{Rafforzamento}(I, C) = \sum_{A \in C} \Delta\text{Forza\_epistemica}(A, I)$$

   dove $\Delta\text{Forza\_epistemica}(A, I) > 0$ quando I fornisce evidenza addizionale per A

3. **Contraddizione e Revisione**: Eliminazione o indebolimento di assunzioni precedenti:

   $$\text{Contraddizione}(I, C) = |\{A \in C \mid I \vdash \neg A\}|$$

Gli effetti cognitivi totali sono una combinazione pesata:

EffettiCognitivi(I, C) = α·|ImplicazioniContestuali(I, C)| + β·Rafforzamento(I, C) - γ·Contraddizione(I, C)

dove α, β, γ sono parametri che riflettono l'importanza relativa dei diversi tipi di effetti.

Sforzo di Elaborazione

Lo sforzo cognitivo richiesto per derivare I da LF dato C include molteplici componenti:

Sforzo(I, C) = Sforzo_Accesso(I, C) + Sforzo_Recupero_Contesto(I, C) + Sforzo_Inferenziale(I, C)

1. **Sforzo di Accesso**: Costo di accedere a rappresentazioni lessicali, concetti, e assunzioni necessarie

   Sforzo_Accesso(I, C) = ∑_{concetto∈I} CostoAccesso(concetto)

   dove CostoAccesso dipende da:

   - Frequenza del concetto (concetti frequenti: basso costo)
   - Recency di attivazione (recentemente usati: basso costo)
   - Connessioni associative (fortemente connessi al contesto: basso costo)
2. **Sforzo di Recupero Contestuale**: Costo di recuperare assunzioni contestuali rilevanti dalla memoria di lavoro o memoria a lungo termine

   Sforzo_Recupero_Contesto(I, C) = NumeroAssunzioni_Richieste(I, C) × Costo_Medio_Recupero

3. **Sforzo Inferenziale**: Complessità computazionale delle inferenze necessarie per derivare effetti cognitivi

   Sforzo_Inferenziale(I, C) = ∑_{inferenza} Complessità(inferenza)

   dove Complessità può essere quantificata attraverso:

   - Numero di passi deduttivi
   - Profondità di annidamento di inferenze
   - Necessità di reasoning non-monotono o abduttivo

Strategia di Comprensione Guidata da Pertinenza

La Strategia di Comprensione Basata sulla Pertinenza può essere formalizzata come un algoritmo euristico:

```
funzione ComprendiEnunciato(LF, C):
  SpazioInterpretazioni = GeneraInterpretazioni(LF, C)
  OrdineAccessibilità = Ordina(SpazioInterpretazioni, per: Sforzo crescente)

  per I in OrdineAccessibilità:
```

```
    EffettiCognitivi = CalcolaEffetti(I, C)
    Sforzo = CalcolaSforzo(I, C)
    Pertinenza = EffettiCognitivi / Sforzo

    se Pertinenza ≥ SogliaAttesa(LF, C):
       se ConsistenteConPresunzioneOttimalità(I, LF, C):
          ritorna I  // Prima interpretazione soddisfacente

  ritorna Fallimento_Comprensione
```

Questo algoritmo cattura l'euristica di elaborazione umana: non cerchiamo esaustivamente tutte le interpretazioni possibili ma procediamo in ordine di accessibilità crescente, arrestandoci appena troviamo un'interpretazione sufficientemente pertinente.

Architettura Computazionale Integrata

Integrando tutti gli stadi elaborati precedentemente, possiamo rappresentare l'architettura completa della comprensione come una cascata computazionale con feedback:

Comprensione(Enunciato_U) = Φ_pragmatica(Φ_semantica(Φ_sintassi(Φ_lessico(Φ_percezione(U)))), C)

Tuttavia, questa caratterizzazione puramente feed-forward è una semplificazione eccessiva. Le evidenze psicolinguistiche e neurolinguistiche rivelano che l'elaborazione è massivamente interattiva, con flussi di informazione bidirezionali:

- Informazione semantica e pragmatica influenza parsing sintattico (risoluzione di ambiguità strutturali)
- Aspettative lessicali influenzano percezione fonologica (effetti Ganong)
- Contesto pragmatico influenza accesso lessicale (priming semantico)
- Conoscenza del mondo influenza interpretazione semantica

Per modellare questa interattività, dobbiamo abbandonare l'architettura feed-forward semplice a favore di un modello dinamico in cui ciascun livello rappresentazionale può influenzare bidirezionalmente gli altri livelli.

Modello Bayesiano Unificato

Una formalizzazione possibile utilizza il framework delle reti bayesiane dinamiche. Sia R = {R_percezione, R_lessico, R_sintassi, R_semantica, R_pragmatica} l'insieme dei livelli rappresentazionali, dove ciascun R_i è una variabile casuale (o vettore di variabili) rappresentante lo stato di quel livello.

Dato un input osservato O = Φ_percezione(U), il problema della comprensione diventa un problema di inferenza bayesiana: trovare la distribuzione di probabilità congiunta posteriore P(R|O) su tutti i livelli dato l'input.

Attraverso il teorema di Bayes e una struttura di dipendenze specificata da un grafo diretto G:

$P(R|O) \propto P(O|R) \cdot P(R) = P(O|R\_percezione) \cdot \prod_i P(R_i \mid Genitori\_G(R_i))$

dove $Genitori\_G(R_i)$ denota l'insieme dei nodi genitori di R_i nel grafo G.

L'interpretazione ottimale corrisponde alla configurazione che massimizza questa probabilità posteriore:

$R^* = \text{argmax}_R P(R|O)$

Questa può essere trovata attraverso algoritmi di inferenza bayesiana come belief propagation o variational inference.

Crucialmente, questo framework permette di modellare naturalmente l'influenza bidirezionale tra livelli. Se per esempio la variabile semantica R_semantica è un genitore della variabile lessicale R_lessico nel grafo G (feedback top-down), allora aspettative semantiche possono modulare l'accesso lessicale attraverso il termine:

$P(R\_lessico \mid R\_semantica)$

nella decomposizione della distribuzione congiunta.

Dinamiche Temporali e Aggiornamento Incrementale

L'elaborazione procede incrementalmente nel tempo. Possiamo modellare questo introducendo una dimensione temporale esplicita. Sia R(t) lo stato del sistema rappresentazionale al tempo t, dove t avanza discretamente man mano che ciascuna nuova parola o chunk di informazione arriva.

L'aggiornamento incrementale è governato da:

$P(R(t+1) \mid O(1:t+1)) \propto P(O(t+1) \mid R(t+1)) \cdot \int P(R(t+1) \mid R(t)) \cdot P(R(t) \mid O(1:t))\, dR(t)$

Questa equazione di filtering bayesiano specifica come la distribuzione posteriore sulle rappresentazioni è aggiornata quando una nuova osservazione O(t+1) arriva, integrando sul

la distribuzione posteriore al tempo precedente $P(R(t)|O(1:t))$ e applicando un modello di transizione $P(R(t+1)|R(t))$ che cattura come le rappresentazioni evolvono, più la verosimiglianza della nuova osservazione $P(O(t+1)|R(t+1))$.

Questo framework cattura:

- **Incrementalità**: Elaborazione parola-per-parola
- **Persistenza**: Informazione da parole precedenti persiste attraverso $P(R(t)|O(1:t))$
- **Predizione**: Il modello di transizione permette predizioni circa stati futuri
- **Revisione**: Quando nuova evidenza contraddice credenze precedenti, la distribuzione posteriore viene aggiornata (garden-path, rianalisi)

Implementazione Neurale e Substrati Biologici della Comprensione

Dal Livello Computazionale all'Implementazione Neurale

Una questione teorica profonda emerge quando consideriamo l'implementazione neurale di questa architettura computazionale astratta. Come possono le operazioni formali che abbiamo specificato essere realizzate attraverso circuiti neurali biologici costituiti da popolazioni di neuroni che comunicano attraverso spike elettrici?

Questo è il problema del livello di implementazione nell'architettura tripartita di Marr: abbiamo caratterizzato il livello computazionale (cosa viene computato e perché) e il livello algoritmico (come viene computato in termini astratti), ma dobbiamo connettere queste caratterizzazioni al livello implementazionale della neurobiologia.

Le evidenze neurolinguistiche forniscono vincoli cruciali. Regioni specifiche della corteccia cerebrale, particolarmente nell'emisfero sinistro, sono selettivamente coinvolte in aspetti diversi dell'elaborazione linguistica:

Mappa Neuroanatomica delle Funzioni Linguistiche

1. **Corteccia Temporale Superiore Posteriore (pSTG)**
    - Funzione: Mappatura suono-significato, accesso lessicale
    - Processo: Conversione rappresentazioni fonologiche → rappresentazioni semantiche
    - Attivazione: Picco 200-400 ms dopo onset stimolo acustico
2. **Giro Frontale Inferiore Sinistro (LIFG)**
    - Pars opercularis (BA44): Elaborazione sintattica, combinazione strutturale
    - Pars triangularis (BA45): Selezione lessicale con vincoli grammaticali
    - Processo: Costruzione strutture gerarchiche, controllo cognitivo linguistico
    - Attivazione: Modulata da complessità sintattica
3. **Giro Angolare (AG)**
    - Funzione: Mantenimento dipendenze sintattiche a lunga distanza
    - Processo: Memoria di lavoro sintattica, gap-filling
    - Attivazione: Aumentata per frasi con movimento sintattico
4. **Giro Temporale Medio (MTG)**
    - Funzione: Integrazione sintassi-semantica, assegnamento ruoli tematici
    - Processo: Mappatura argomenti a ruoli (agente, paziente, tema)
    - Lesioni: Deficit nella comprensione di frasi reversibili
5. **Reti Semantiche Distribuite**
    - Corteccia prefrontale ventrale e laterale: Controllo semantico, selezione
    - Poli temporali: Hub semantico transmodale

- Giro fusiforme: Elaborazione semantica visiva (parole scritte)

Dinamica Temporale e Oscillazioni Neurali

Oltre alla localizzazione anatomica, la dinamica temporale rivelata attraverso EEG e MEG mostra che queste regioni interagiscono attraverso oscillazioni neurali sincronizzate in diverse bande di frequenza.

Bande di Frequenza e Loro Funzioni Computazionali

Oscillazioni Delta (1-4 Hz): Tracking Prosodico

Funzione: Segmentazione di unità linguistiche grandi (frasi, enunciati)

Le oscillazioni delta si agganciano (entrain) al ritmo prosodico del discorso. Picchi delta corrispondono a confini frasali, permettendo predizione temporale di quando informazione critica arriverà.

Formalmente, la fase delta al tempo t, $\varphi_\delta(t)$, traccia la struttura prosodica:

$\varphi_\delta(t) \propto$ StrutturaProsodica(Discorso(t))

Quando $\varphi_\delta \approx$ picco → Aspettativa di confine frasale → Preparazione per integrazione

Evidenza empirica:

- Coerenza δ tra discorso e attività neurale: $r > 0,6$
- Disruzione δ (attraverso stimolazione transcranica) → Difficoltà segmentazione frasale

Oscillazioni Theta (4-8 Hz): Chunking e Working Memory

Funzione: Organizzazione temporale di chunks lessicali e sintattici, supporto memoria di lavoro

Ciascun ciclo theta (~150-200 ms) corrisponde approssimativamente alla durata di elaborazione di un chunk (sillaba, morfema, parola breve).

La capacità di working memory è limitata dal numero di items che possono essere mantenuti in cicli theta distinti:

$\text{Capacità\_WM} \approx \text{Frequenza\_}\theta \cdot \text{Durata\_Mantenimento} / \text{Durata\_Ciclo\_}\theta \approx 5\ \text{Hz} \cdot 2\ \text{s} / 0,2\ \text{s} \approx 7\pm2$ items (limite di Miller)

Evidenza:

- Potenza θ aumenta durante elaborazione di frasi sintatticamente complesse
- Potenza θ correlata con performance in task di memoria sintattica: $r = 0,72$

Oscillazioni Alpha (8-13 Hz): Inibizione e Gating Attenzionale

Funzione: Soppressione di informazione irrilevante, allocazione attenzionale

Alpha sincronizzazione (↑ potenza) in regioni che devono essere inibite Alpha desincronizzazione (↓ potenza) in regioni attivamente processanti

Durante ascolto linguistico:

- Corteccia uditiva: Alpha ↓ (desincronizzazione - attivo)
- Corteccia visiva: Alpha ↑ (sincronizzazione - inibito)

Meccanismo: Alpha ritmica implementa gating dell'informazione. La fase alpha determina l'eccitabilità neuronale:

Eccitabilità(Regione, t) ∝ cos(Fase_α(Regione, t) - π)

Massima eccitabilità quando fase ≈ π (picco dell'onda)

Oscillazioni Beta (13-30 Hz): Feedback Top-Down e Mantenimento

Funzione: Comunicazione feedback da corteccia frontale a regioni posteriori, implementazione aspettative top-down

Beta accoppiamento tra regioni frontali e temporali durante predizione linguistica. Beta aumenta prima di parole altamente predicibili.

Meccanismo: Beta implementa segnali di predizione che modulano processamento bottom-up:

Predizione_Top-Down(t) = β_Frontale(t) → Modulazione_Temporale(t+Δt)

Quando parola predetta arriva:

- Mismatch piccolo → Beta mantiene
- Mismatch grande → Beta reset → Burst gamma per risoluzione

Oscillazioni Gamma (30-100 Hz): Binding Locale

Funzione: Binding di features in rappresentazioni coerenti, comunicazione feedforward rapida

Neuroni codificanti features correlate scaricano sincronicamente in gamma. Questa sincronizzazione binda features in oggetto/concetto unificato.

Rappresentazione di un concetto C:

Rappresentazione(C) = {Neurone_i | Fase_γ(Neurone_i) ≈ Fase_γ(Neurone_j) ∀i,j codificanti features di C}

Il problema del binding è risolto attraverso temporal synchrony:

- Concetto_1: Neuroni $\{n_1, n_2, n_3\}$ sincronizzati in gamma fase φ_1

- Concetto_2: Neuroni $\{n_4, n_5, n_6\}$ sincronizzati in gamma fase φ_2
- Dove $\varphi_1 \neq \varphi_2$ (gamma asincrona) → concetti distinti mantenuti separati

Evidenza:

- Burst gamma durante accesso lessicale (200-300 ms post-onset)
- Sincronizzazione gamma tra regioni che devono comunicare rapidamente
- Gamma aumenta per violazioni semantiche (N400) - riflette risoluzione conflitto

Implementazione Neurale del Binding Composizionale

Come possiamo connettere questi substrati neurali alle operazioni computazionali formali di composizione semantica? Una proposta teorica influente è che le operazioni di binding composizionale sono implementate attraverso sincronizzazione temporale di popolazioni neuronali.

Binding Composizionale via Sincronizzazione Temporale

Per la frase "Il gatto insegue il topo", la rappresentazione semantica compositiva richiede di mantenere distinte due strutture di ruoli tematici:

- Struttura 1: GATTO come AGENTE di INSEGUIRE
- Struttura 2: TOPO come PAZIENTE di INSEGUIRE

Implementazione neurale:

Assembly_1 = {Neuroni_GATTO ∪ Neuroni_AGENTE} sincronizzati con fase γ_1

Assembly_2 = {Neuroni_TOPO ∪ Neuroni_PAZIENTE} sincronizzati con fase γ_2

dove $\varphi\_\gamma_1 \neq \varphi\_\gamma_2$ mantiene le strutture separate.

Formalmente, un pattern di attivazione neuronale distribuito $R(t) = \{r_1(t), r_2(t), ..., r\_N(t)\}$ dove $r\_i(t)$ è il tasso di firing del neurone i, implementa binding composizionale attraverso vincoli sulla coerenza di fase:

Per tutti i neuroni $i, j \in$ Assembly_k: $Fase\_\gamma(r\_i(t)) \approx Fase\_\gamma(r\_j(t))$

Per neuroni in assemblies diverse A_k e A_l: $Fase\_\gamma(r\_i \in A\_k) \neq Fase\_\gamma(r\_j \in A\_l)$

La fase gamma di un segnale neuronale r(t) può essere estratta attraverso:

$Fase\_\gamma(r(t)) = arg(H(BandPass\_\gamma(r(t))))$

dove H è la trasformata di Hilbert e BandPass_γ filtra il segnale nella banda gamma 30-100 Hz.

Limiti di Capacità e Memoria di Lavoro Sintattica

Questa implementazione fornisce una spiegazione naturale per i limiti di capacità nella comprensione di strutture sintattiche complesse.

Vincolo di Capacità da Separazione di Fase

Il sistema può mantenere soltanto M assemblee simultaneamente attive con fasi gamma sufficientemente distinte prima che interferenza emerga:

$M = \lfloor 2\pi / \Delta\varphi_{min} \rfloor$

dove $\Delta\varphi_{min}$ è la distinzione di fase minima richiesta per evitare confusione.

Basandosi su evidenze neurofisiologiche, $\Delta\varphi_{min} \approx \pi/4$ radianti, quindi:

$M \approx \lfloor 2\pi / (\pi/4) \rfloor = 8$ assemblee

Questo corrisponde al limite classico di 7±2 items per la capacità della memoria di lavoro (Miller, 1956), suggerendo che il limite di Miller riflette vincoli fondamentali dell'architettura neurale di binding attraverso sincronizzazione temporale.

Applicazione alla Comprensione di Frasi Annidate

Frasi con annidamento sintattico profondo richiedono il mantenimento simultaneo di molteplici strutture sintattiche incomplete:

"Il topo che il gatto che il cane inseguiva vedeva scappò"

Strutture da mantenere simultaneamente:

1. [Il topo ... scappò] - clausola principale incompleta
2. [che il gatto ... vedeva] - relativa 1 incompleta
3. [che il cane inseguiva] - relativa 2 completa

Ciascuna struttura incompleta richiede un'assembly dedicata. Con limite $M \approx 8$, annidamenti oltre 2-3 livelli causano difficoltà severe, consistente con evidenze comportamentali.

Accoppiamento Cross-Frequenza

Le oscillazioni in diverse bande interagiscono attraverso accoppiamenti cross-frequenza, implementando relazioni gerarchiche tra livelli di elaborazione.

Accoppiamento Theta-Gamma: Chunking Gerarchico

La fase theta organizza *quando* chunks sono processati, mentre l'ampiezza gamma codifica *cosa* (contenuto del chunk).

Formalmente, se $A_\gamma(t)$ è l'ampiezza gamma e $\varphi_\theta(t)$ è la fase theta:

$A_\gamma(t) = A_{baseline} + A_{modulazione} \cdot \cos(\varphi_\theta(t) - \varphi_{preferita})$

dove $\varphi_{preferita}$ è la fase theta ottimale per burst gamma, tipicamente il picco.

Questo accoppiamento implementa organizzazione gerarchica:

- Chunks di livello inferiore (morfemi, parole) → elaborati in burst gamma individuali
- Chunks di livello superiore (sintagmi, clausole) → sequenze di cicli theta
- Relazione gerarchica: Posizione nel ciclo theta specifica ruolo strutturale

Accoppiamento Delta-Theta: Multi-Scala Temporale

Similmente, la fase delta può modulare l'ampiezza theta, organizzando chunks theta in unità ancora più grandi:

$A_\theta(t) = A_{baseline,\theta} + A_{modulazione,\theta} \cdot \cos(\varphi_\delta(t) - \varphi_{preferita,\delta})$

Questa cascata di accoppiamenti implementa gerarchia temporale multi-scala:

Delta (1-4 Hz) → Frasi, enunciati (250-1000 ms) ↓ modula Theta (4-8 Hz) → Sintagmi, parole (125-250 ms) ↓ modula
Gamma (30-100 Hz) → Morfemi, sillabe (10-30 ms)

Modello Dinamico Formale di Oscillatori Accoppiati

Possiamo formalizzare questa architettura oscillatoria come un sistema dinamico di oscillatori accoppiati operanti a scale temporali multiple.

Sia $x_\delta(t)$, $x_\theta(t)$, $x_\gamma(t)$ le variabili di fase delle oscillazioni delta, theta, gamma in una regione corticale R.

L'evoluzione temporale è governata da equazioni differenziali accoppiate:

Equazioni di Evoluzione

$$dx_\delta/dt = \omega_\delta + \sum_j K_{\delta j} \cdot \sin(x_\delta^j(t) - x_\delta(t)) + I_\delta(t)$$

$$dx_\theta/dt = \omega_\theta + \sum_j K_{\theta j} \cdot \sin(x_\theta^j(t) - x_\theta(t)) + \alpha \cdot \cos(x_\delta(t)) + I_\theta(t)$$

$$dx_\gamma/dt = \omega_\gamma + \sum_j K_{\gamma j} \cdot \sin(x_\gamma^j(t) - x_\gamma(t)) + \beta \cdot \cos(x_\theta(t)) + I_\gamma(t)$$

dove:

- ω_δ, ω_θ, ω_γ: frequenze naturali intrinseche
- $K_{\delta j}$, $K_{\theta j}$, $K_{\gamma j}$: forze di accoppiamento tra oscillatori nella stessa banda in regioni diverse
- α, β: parametri di accoppiamento cross-frequenza (delta→theta, theta→gamma)
- $I_\delta(t)$, $I_\theta(t)$, $I_\gamma(t)$: input esterni (stimolazione linguistica, feedback da altre regioni)

Questo sistema cattura:

- **Sincronizzazione intra-banda**: Termine $\sum K_{ij} \sin(...)$ causa oscillatori nella stessa banda a sincronizzarsi

- **Modulazione cross-banda**: Termini α cos(x_δ) e β cos(x_θ) implementano modulazione di ampiezza
- **Driving esterno**: I(t) permette influenza da input linguistico e feedback

Meccanismo di Segmentazione via Reset di Fase

Le oscillazioni forniscono un meccanismo naturale per segmentazione dell'input continuo. Se oscillazioni theta si agganciano a unità lessicali ricorrenti, dopo sufficiente esposizione la fase theta diventa predittiva dei confini lessicali.

Apprendimento di Segmentazione

Il discorso arriva come flusso continuo: /ilgattoèsultavolo/

Tracciamento di probabilità transizionali (TPs):

TP(B|A) = P(B segue A) = Count(AB) / Count(A)

TPs alte all'interno di parole, TPs basse ai confini.

Implementazione oscillatoria:

Quando TP alta → Nessun reset di fase theta (continuità all'interno della parola) Quando TP bassa → Reset di fase theta (confine di parola)

Formalmente, il reset è governato da:

Se TP(sillaba_{i+1} | sillaba_i) < θ_reset: φ_θ(t) → 0 (reset di fase) Altrimenti: φ_θ(t) → φ_θ(t) + ω_θ·Δt (evoluzione continua)

Dopo apprendimento, la struttura di reset di fase codifica la segmentazione lessicale, permettendo parsing rapido di flusso continuo in unità discrete.

Predizioni Testabili del Modello Oscillatorio

Questo modello genera predizioni empiriche specifiche testabili attraverso MEG/EEG:

Predizione 1: Accoppiamento Theta-Gamma Aumenta con Complessità

Frasi sintatticamente complesse (annidamento multiplo) dovrebbero mostrare:

- Accoppiamento theta-gamma più forte
- Più assemblee gamma distinte mantenute nel ciclo theta

Misura: Phase-Amplitude Coupling (PAC) theta-gamma Test: PAC_complesso > PAC_semplice

Predizione 2: Reset Delta a Confini Frasali

Confini prosodici maggiori dovrebbero causare reset di fase delta.

Misura: Inter-Trial Phase Coherence (ITPC) delta allineata a confini frasali Test: ITPC_confini > ITPC_dentro-frase

Predizione 3: Separazione di Fase Gamma per Concetti Distinti

Processamento simultaneo di concetti distinti dovrebbe produrre burst gamma con fasi diverse.

Misura: Differenza di fase γ tra popolazioni neuronali codificanti concetti diversi Test: $|\Delta\varphi_\gamma| \approx \pi$ (massima separazione) per concetti simultanei distinti

Predizione 4: Disruzione Oscillatoria → Deficit Specifici

Stimolazione transcranica ritmica (rhythmic TMS) a frequenze specifiche dovrebbe causare deficit selettivi:

- Disruzione theta (6 Hz) → Deficit memoria di lavoro sintattica
- Disruption gamma (40 Hz) → Deficit binding semantico
- Disruzione delta (2 Hz) → Deficit segmentazione prosodica

Queste predizioni collegano direttamente il modello teorico a fenomenologia osservabile, permettendo validazione o falsificazione empirica.

Semantica Grounded e Rappresentazioni Multi-Modali

Oltre la Semantica Simbolica Astratta

Un'altra area in cui l'integrazione teorica rimane incompleta è la relazione tra significato linguistico e significato non-linguistico, particolarmente il significato di rappresentazioni percettive e motorie.

La semantica formale tradizionale tratta il significato linguistico come una relazione astratta tra espressioni e entità in modelli logici, dove le entità sono oggetti matematici come insiemi, funzioni, e mondi possibili. Questa concezione ha il pregio della precisione formale ma sembra disconnessa dalla realtà psicologica di come gli umani effettivamente rappresentano e comprendono significati.

Teoria della Cognizione Incarnata (Embodied Cognition)

I teorici della cognizione incarnata e grounded cognition argomentano che i significati concettuali non sono simboli astratti arbitrariamente correlati con referenti esterni ma piuttosto sono costituiti da o strettamente connessi a rappresentazioni sensori-motorie derivate dall'esperienza percettiva e interattiva con il mondo.

Secondo questa prospettiva, comprendere la parola "tazza" non consiste semplicemente nell'accedere a una rappresentazione astratta del concetto TAZZA ma coinvolge la parziale riattivazione dei sistemi neurali sensoriali e motori che sono stati attivati durante esperienze effettive con tazze:

- Corteccia visiva: Codifica forma tipica di una tazza
- Corteccia somatosensoriale e motoria: Codifica azione di afferrare il manico
- Corteccia gustativa e olfattiva: Codifica esperienza di bere
- Corteccia insulare: Codifica sensazioni tattili (calore, peso)

Evidenze Empiriche per Semantica Grounded

1. Attivazione Somatotopica per Verbi di Azione

Gli studi di neuroimaging funzionale (fMRI) rivelano che la comprensione di verbi che denotano azioni specifiche attiva selettivamente regioni della corteccia motoria corrispondenti agli effettori corporei coinvolti:

Studio seminale (Hauk, Johnsrude & Pulvermüller, 2004):

- Stimoli: Verbi di azione con piede ("calciare"), mano ("afferrare"), viso ("leccare")
- Risultato fMRI:
    - "calciare" → Attivazione corteccia motoria piede (superiore)
    - "afferrare" → Attivazione corteccia motoria mano (laterale)
    - "leccare" → Attivazione corteccia motoria viso/bocca (inferiore)

Organizzazione somatotopica preservata: La mappa corporea nella corteccia motoria è riattivata durante la comprensione linguistica.

Formalmente, possiamo rappresentare questo come:

Comprensione("calciare") → Attivazione(CorteMotoria_Piede) > baseline

dove CorteMotoria_Piede è localizzata topograficamente secondo homunculus motorio.

2. Attivazione Retinotopica per Parole Visive

Parole che denotano oggetti con proprietà visive distintive attivano regioni della corteccia visiva selettive per quelle proprietà.

Studio (Pulvermüller & Hauk, 2006):

- Stimoli: Parole per oggetti verticali ("torre", "pino") vs oggetti orizzontali ("fiume", "serpente")
- Risultato MEG:
    - Parole verticali → Attivazione dorsale in corteccia visiva (processa orientamento verticale)
    - Parole orizzontali → Attivazione ventrale in corteccia visiva (processa orientamento orizzontale)

Latenza: 150-200 ms post-onset parola (precoce, automatica)

3. Compatibilità Azione-Frase (ACE - Action-sentence Compatibility Effect)

Comprendere frasi che descrivono azioni in certe direzioni facilita risposte motorie nella stessa direzione.

Paradigma (Glenberg & Kaschak, 2002):

- Task: Giudicare se frase ha senso
- Frasi: "Chiudi il cassetto" (movimento verso sé) vs "Apri il cassetto" (movimento da sé)
- Risposta: Movimento verso/da sé per indicare sì/no

Risultato:

- Facilitazione quando direzione frase = direzione risposta
- Interferenza quando direzione frase ≠ direzione risposta

RT("Chiudi", risposta-verso) < RT("Chiudi", risposta-da) RT("Apri", risposta-da) < RT("Apri", risposta-verso)

Interpretazione: Comprensione attiva automaticamente rappresentazioni motorie direzionalmente specifiche che facilitano o interferiscono con esecuzione motoria.

4. Switch Percettivo Orientazionale

Dopo comprendere frase che descrive oggetto in orientamento specifico, riconoscimento visivo è facilitato per immagini nello stesso orientamento.

Studio (Stanfield & Zwaan, 2001):

- Fase 1: Leggere frase
    - "Giovanni ha piantato il chiodo nella parete" (chiodo orizzontale)
    - "Giovanni ha piantato il chiodo nel pavimento" (chiodo verticale)
- Fase 2: Verificare se immagine corrisponde a oggetto nella frase
    - Immagine chiodo orizzontale vs verticale

Risultato:

- RT più rapido quando orientamento immagine = orientamento implicato da frase
- RT(match) = 820 ms < RT(mismatch) = 910 ms, $p < 0,01$

Latenza effetto: 200-400 ms → Emerge durante comprensione, non post-comprensione

Formalizzazione: Rappresentazioni Semantiche Multi-Componenziali

Integrare questa prospettiva grounded con la semantica formale richiede di arricchire i nostri modelli semantici per includere non soltanto relazioni astratte con entità in modelli logici ma anche connessioni con rappresentazioni percettive e motorie.

Modello Multi-Modale del Significato Lessicale

Il significato di una parola lessicale è una tupla multi-componenziale:

$M = \langle S, P, V, A, E, ... \rangle$

dove:

- **S**: Componente simbolica astratta (denotazione in modello logico)
- **P**: Rappresentazione percettiva (features visive, uditive, tattili)
- **V**: Rappresentazione motoria (azioni, affordances)
- **A**: Rappresentazione affettiva (valenza emotiva, arousal)
- **E**: Rappresentazione episodica (memoria di eventi specifici)

Componente Simbolica S

La componente tradizionale della semantica formale:

$S \in$ Modello_Logico

Per esempio, per "tazza": $S = \lambda x$ [Tazza(x)] $\in \langle e,t \rangle$

Questa componente supporta ragionamento inferenziale:

- Se Tazza(x) allora Contenitore(x) (inferenza tassonomica)
- Se Tazza(x) allora PuòContenere(x, Liquido) (inferenza funzionale)

Componente Percettiva P

Rappresentazione distribuita in spazi di features percettive ad alta dimensionalità:

$P = \langle p_1, p_2, ..., p\_D \rangle \in \mathbb{R}^D$

dove ciascuna dimensione corrisponde a una feature percettiva:

- Visuale: Orientamento bordi, colore (HSV), texture, forma globale, ...
- Uditiva: Frequenza, intensità, timbre, pattern temporali, ...
- Tattile: Rugosità, temperatura, peso, solidità, ...

Per "tazza":

P_visuale = [orientamento_verticale: 0,8, forma_cilindrica: 0,9, colore_variabile: 0,5, presenza_manico: 0,95, ...]

P_tattile = [temperatura_calda: 0,7, superficie_liscia: 0,8, peso_medio: 0,6, ...]

Questo vettore rappresenta il pattern di attivazione tipico prodotto nelle cortecce sensoriali quando il concetto TAZZA è attivato.

Componente Motoria V

Rappresentazione di azioni e affordances associate:

$V = \{\langle azione_1, probabilità_1 \rangle, \langle azione_2, probabilità_2 \rangle, ...\}$

Per "tazza":

V = { ⟨afferrare_manico, 0,95⟩, ⟨portare_a_bocca, 0,85⟩, ⟨riempire, 0,7⟩, ⟨appoggiare, 0,8⟩, ... }

Ciascuna azione è associata con pattern di attivazione nella corteccia motoria primaria e premotoria che specifica i programmi motori.

Componente Affettiva A

Rappresentazione della valenza emotiva e arousal:

A = ⟨valenza, arousal, dominanza⟩ ∈ $\mathbb{R}^3$

Misurata tipicamente attraverso Self-Assessment Manikin (SAM) o Affective Norms:

Per "tazza": A = ⟨valenza: +0,3, arousal: 0,1, dominanza: 0,5⟩ (leggermente positivo, basso arousal, neutrale dominanza)

Per "ragno": A = ⟨valenza: -0,6, arousal: 0,7, dominanza: -0,4⟩
(negativo, alto arousal per molte persone, bassa dominanza)

Composizione Semantica Multi-Modale

Quando combiniamo parole in frasi, dobbiamo comporre non soltanto le componenti simboliche ma anche le componenti grounded.

Composizione di Componenti Simboliche

Standard nella semantica formale, attraverso applicazione funzionale o altre operazioni compositive:

S("mela rossa") = S("rossa")(S("mela")) = λx[Rosso(x) ∧ Mela(x)]

Composizione di Componenti Percettive

La rappresentazione percettiva del sintagma è costruita modulando la rappresentazione del nome con le features dell'aggettivo:

P("mela rossa") = P("mela") + α · ΔP_colore("rosso")

dove:

- P("mela") è il vettore percettivo base
- ΔP_colore("rosso") è un vettore di modificazione che enfatizza features di colore rosso
- α è un parametro di forza di modulazione (0 < α < 1)

Matematicamente, attraverso addizione pesata:

P_risultato[dimensione_i] = P_base[dimensione_i] + α · Δ[dimensione_i]

Per features di colore: P_risultato[rosso] = P("mela")[rosso] + α · 1,0 = aumentato P_risultato[verde] = P("mela")[verde] + α · (-0,5) = diminuito

Composizione di Componenti Motorie

Per costruzioni verbo-oggetto, la componente motoria del verbo si combina con le affordances dell'oggetto:

V("lanciare palla") = Unifica(V("lanciare"), V("palla"))

dove Unifica seleziona azioni compatibili:

- Da V("lanciare"): {movimento_braccio, rilascio, ...}
- Da V("palla"): {afferrare_sferica, rotolare, rimbalzare, ...}
- Unificato: {afferrare_sferica → movimento_braccio → rilascio → volo_parabolico}

Questa sequenza motoria è una simulazione mentale dell'azione complessa.

Costruzione di Simulazioni Dinamiche

Per frasi intere che descrivono eventi, la rappresentazione grounded corrisponde a una simulazione mentale dinamica:

"Il gatto insegue il topo"

Componenti della simulazione:

1. **Visuale dinamica**:
    - Rappresentazioni visive di gatto e topo in movimento relativo
    - Traiettorie spaziali (topo fugge, gatto segue)
    - Features cinematiche (velocità, accelerazione)
2. **Motoria**:
    - Pattern motori di inseguimento (se simulato dal punto di vista del gatto)
    - O pattern motori di fuga (se simulato dal punto di vista del topo)
3. **Affettiva**:
    - Eccitazione, tensione dell'inseguimento
    - Paura (dal punto di vista del topo)
    - Eccitazione predatoria (dal punto di vista del gatto)

Formalmente, la simulazione è una traiettoria nello spazio di stati percettivi-motori:

Simulazione(t) = {P(t), V(t), A(t)} per t ∈ [0, T]

dove t varia temporalmente durante la durata dell'evento simulato.

Evidenze Comportamentali per Simulazioni durante Comprensione

Gli effetti di compatibilità azione-frase che abbiamo menzionato costituiscono evidenza che queste simulazioni grounded sono parte integrante della comprensione, non epifenomeni post-comprensione.

Compatibilità Direzionale

Studio esteso (Glenberg & Kaschak, 2002):

- Molteplici frasi testate con azioni implicite verso/da sé
- Effetto robusto: RT_compatibile < RT_incompatibile
- Differenza media: ~50 ms (significativo)
- Latenza emergenza: 200-400 ms post-onset frase critica

Interpretazione: La comprensione attiva rappresentazioni motorie direzionalmente specifiche che facilitano o interferiscono con esecuzione motoria concorrente.

Affordances Visive dopo Comprensione

Studio (Zwaan & Taylor, 2006):

- Fase 1: Leggere frase descrivente interazione con oggetto
    - "Giovanni ha girato la manopola" (rotazione oraria)
- Fase 2: Giudicare orientamento immagine di manopola
    - Immagine pre-ruotata in varie direzioni

Risultato: Facilitazione quando orientamento immagine = orientamento risultante dall'azione descritta

Interpretazione: La simulazione mentale dell'azione produce rappresentazione visiva dello stato risultante, facilitando riconoscimento di immagini in quell'orientamento.

Integrazione Simbolico-Percettivo: Architettura Ibrida

Una teoria unificata deve integrare le componenti simboliche (necessarie per ragionamento logico e composizionalità) con le componenti grounded (necessarie per ancoraggio percettivo e risonanza embodied).

Modello a Due Vie (Dual-Coding)

Basandosi sulla teoria dual-coding di Paivio, possiamo postulare due sistemi rappresentazionali parzialmente indipendenti ma interconnessi:

1. **Sistema Verbale-Simbolico**
    - Rappresentazioni proposizionali astratte
    - Struttura logica compositiva
    - Supporta ragionamento inferenziale formale
    - Implementato in reti linguistiche (LIFG, MTG, ATL)
2. **Sistema Immaginativo-Percettivo**

- Rappresentazioni modali distribuite
- Formato analogico (preserva topologia spaziale, temporale)
- Supporta simulazioni mentali
- Implementato in cortecce sensoriali e motorie

Connessioni Referenziali tra Sistemi

I due sistemi sono connessi attraverso links referenziali:

Parola_w ←→ Concetto_Verbale_w ←→ Rappresentazione_Percettiva_w

L'attivazione di una parola attiva sia la rappresentazione verbale sia (attraverso links referenziali) la rappresentazione percettiva corrispondente.

Formalmente, l'attivazione si propaga:

Attivazione_totale(Concetto_C) = Attivazione_verbale(C) + β · Attivazione_percettiva(C)

dove β è un parametro che modula quanto fortemente le rappresentazioni percettive contribuiscono all'attivazione totale.

Context-Dependent Weighting

Il peso relativo dei due sistemi varia con il task e il contesto:

β = f(Task, Contesto)

- Task astratto (matematica, logica): β basso (predominanza verbale-simbolica)
- Task concreto (descrizione scena, navigazione): β alto (predominanza percettiva)
- Contesto ricco di dettagli percettivi: β alto
- Contesto astratto: β basso

Evidenza: Neuroimaging mostra attivazione differenziale:

- Frasi astratte ("La verità è relativa"): Predominanza reti verbali (LIFG, ATL)
- Frasi concrete ("Il gatto insegue il topo"): Attivazione motoria e visiva addizionale

Similarità Semantica Multi-Modale

La similarità semantica tra concetti può essere quantificata combinando similarità nelle diverse componenti:

$Similarità\_totale(C_1, C_2) = w\_S \cdot Sim\_S(C_1, C_2) + w\_P \cdot Sim\_P(C_1, C_2) + w\_V \cdot Sim\_V(C_1, C_2) + w\_A \cdot Sim\_A(C_1, C_2)$

dove w_S + w_P + w_V + w_A = 1 sono pesi normalizzati.

Similarità Simbolica

Basata su sovrapposizione di denotazioni o features semantiche discrete:

$Sim\_S(C_1, C_2) = |S_1 \cap S_2| / |S_1 \cup S_2|$ (Jaccard similarity)

Per concetti tassonomici: Sim_S("cane", "gatto") = alta (entrambi animali, mammiferi, domestici) Sim_S("cane", "tavolo") = bassa (categorie ontologiche distanti)

Similarità Percettiva

Basata su distanza nello spazio percettivo:

$Sim\_P(C_1, C_2) = \cos(P_1, P_2) = (P_1 \cdot P_2) / (||P_1|| \cdot ||P_2||)$

Dove P_1, P_2 sono vettori di features percettive.

Per features visive: Sim_P("arancia", "palla") = alta (entrambi rotondi, simile dimensione) Sim_P("arancia", "banana") = media (forme diverse, colori simili)

Similarità Motoria

Basata su sovrapposizione di affordances:

$Sim\_V(C_1, C_2) = |V_1 \cap V_2| / |V_1 \cup V_2|$

dove V sono insiemi di azioni possibili.

Sim_V("tazza", "bicchiere") = alta (afferrare, bere, riempire simili) Sim_V("tazza", "piatto") = bassa (azioni diverse)

Similarità Affettiva

Basata su distanza nello spazio affettivo tridimensionale:

$Sim\_A(C_1, C_2) = 1 / (1 + ||A_1 - A_2||)$

dove $||A_1 - A_2||$ è distanza euclidea tra vettori affettivi.

Sim_A("cucciolo", "fiore") = alta (entrambi positivi, basso arousal) Sim_A("cucciolo", "serpente") = bassa (valenze opposte per molte persone)

Predizioni del Modello Multi-Modale

Questo modello integrato genera predizioni testabili:

Predizione 1: Doppia Dissociazione

Lesioni a sistemi verbali vs percettivi dovrebbero causare deficit selettivi:

- Lesione LIFG/ATL: Deficit elaborazione simbolica (giudizi tassonomici, inferenze logiche) ma preservazione giudizi percettivi (similarità visiva)
- Lesione cortecce sensoriali: Deficit rappresentazioni percettive ma preservazione elaborazione simbolica

Evidenza: Pazienti con afasia semantica (lesioni ATL) sono compromessi in task verbali ma preservati in matching visivo di oggetti.

Predizione 2: Effetti di Concretezza

Parole concrete (alto imageability) dovrebbero mostrare:

- Attivazione più ampia (verbale + percettiva)
- Processing più rapido (doppia codifica)
- Memoria migliore (dual-trace)

Evidenza: Effetto concretezza robusto in memoria: recall(concreto) > recall(astratto)

Predizione 3: Context-Dependent Activation

L'attivazione di componenti percettive dovrebbe variare con richieste del task:

- Task semantico profondo: Attivazione percettiva alta
- Task fonologico superficiale: Attivazione percettiva ridotta

Evidenza fMRI: Livelli di elaborazione modulano attivazione cortecce sensoriali.

Predizione 4: Facilitazione Cross-Modale

Priming dovrebbe funzionare attraverso modalità:

Visual prime (immagine martello) → Facilitazione parola "martello" Motor prime (pantomima martellare) → Facilitazione parola "martello"

Evidenza: Priming cross-modale robusto per concetti concreti.

L'Emergere della Competenza Linguistica: Una Prospettiva di Sviluppo Integrata

Il Periodo Prelinguistico e la Preparazione Cognitiva al Linguaggio

Quando osserviamo un neonato nei suoi primi giorni di vita, potremmo essere tentati di considerarlo un organismo essenzialmente passivo, limitato a risposte riflessive e a manifestazioni vegetative di base. Tuttavia, questa impressione superficiale maschera una realtà profondamente diversa: già nelle prime ore dopo la nascita, il sistema percettivo del neonato dimostra capacità sorprendentemente sofisticate che costituiscono i prerequisiti essenziali per l'acquisizione linguistica. Le ricerche condotte utilizzando il paradigma della suzione ad alta ampiezza, una metodologia ingegnosa che sfrutta il riflesso di suzione innato come indicatore delle preferenze percettive del neonato, hanno rivelato che bambini di pochi giorni di vita sono già capaci di discriminare contrasti fonetici sottili che saranno cruciali per l'apprendimento della loro lingua nativa.

Il meccanismo attraverso cui questa capacità discriminativa viene misurata merita una descrizione dettagliata, poiché illustra l'ingegnosità metodologica della ricerca sullo sviluppo

infantile. Il bambino viene posto in una situazione confortevole con un ciuccio speciale collegato a un trasduttore di pressione che misura con precisione millimetrica l'intensità e la frequenza della suzione. Inizialmente, quando un nuovo stimolo acustico viene presentato ripetutamente, il bambino mostra un aumento dell'attività di suzione, un fenomeno che riflette l'orientamento attenzionale verso la novità. Tuttavia, attraverso presentazioni ripetute dello stesso stimolo, si verifica un processo di abituazione: la frequenza della suzione diminuisce progressivamente, segnalando che il sistema percettivo del neonato ha formato una rappresentazione mentale dello stimolo e lo riconosce come familiare. Questo declino nell'attenzione costituisce la fase critica dell'esperimento. A questo punto, se viene introdotto un nuovo stimolo acustico che il bambino percepisce come differente dal primo, si osserva una disabituazione: la frequenza di suzione aumenta nuovamente in modo significativo, indicando che il sistema percettivo ha rilevato una differenza qualitativa tra i due stimoli.

Utilizzando questa metodologia, i ricercatori hanno documentato che neonati di appena due o tre giorni di vita sono capaci di discriminare coppie di sillabe come "ba" e "pa", che differiscono soltanto nella caratteristica temporale del voice onset time, ovvero l'intervallo tra il rilascio della consonante e l'inizio della fonazione vocalica. Questa distinzione temporale, che nell'adulto corrisponde alla differenza tra consonanti sorde e sonore, avviene su scale temporali dell'ordine di venti o trenta millisecondi, eppure il sistema uditivo neonatale è già sintonizzato su questa dimensione acustica critica. Ancora più sorprendentemente, i neonati dimostrano capacità di discriminazione che vanno oltre i contrasti rilevanti per la loro lingua nativa futura: possono distinguere contrasti fonetici che non sono fonologicamente distintivi nella lingua dell'ambiente circostante, ma che lo sono in altre lingue del mondo.

Questa sensibilità universale ai contrasti fonetici solleva interrogativi profondi sulla natura dell'apparato percettivo innato. Una possibilità è che il sistema uditivo umano sia geneticamente predisposto a una modalità di analisi del segnale acustico che estrae automaticamente dimensioni fisiche particolarmente rilevanti per la comunicazione vocale umana. L'analisi spettrale del segnale sonoro operata dalla coclea e dalle stazioni uditive successive nel tronco encefalico fornisce al cervello una rappresentazione tempo-frequenza del suono in arrivo, nella quale transizioni rapide di energia spettrale, come quelle che caratterizzano le consonanti, costituiscono eventi salienti che attirano automaticamente l'elaborazione neurale. La consonante "b", per esempio, si manifesta come una transizione rapida da uno stato di silenzio a un pattern di energia acustica concentrata in bande di frequenza specifiche, con formanti che variano nel tempo secondo traiettorie caratteristiche. Il sistema percettivo neonatale, pur non avendo ancora esperienza con il linguaggio, è già configurato per rilevare tali discontinuità acustiche e per rappresentarle come eventi discreti.

Tuttavia, questa spiegazione puramente acustica della discriminazione fonetica neonatale si scontra con un fenomeno intrigante scoperto negli anni settanta: la percezione categoriale del linguaggio. Quando adulti ascoltano una serie di stimoli sintetizzati che variano lungo un continuum acustico, per esempio variando gradualmente il voice onset time da valori tipici di "ba" a valori tipici di "pa", non percepiscono un cambiamento graduale della qualità sonora. Invece, esperiscono una transizione brusca: fino a un certo punto critico del continuum, tutti gli stimoli vengono percepiti come "ba", mentre oltre quel punto tutti vengono percepiti come "pa". La funzione di identificazione mostra quindi una transizione a gradino, mentre la funzione di discriminazione mostra un picco pronunciato al confine categoriale: coppie di stimoli fisicamente equidistanti lungo il continuum vengono discriminate molto meglio se

attraversano il confine categoriale rispetto a coppie che giacciono dalla stessa parte del confine. Questo pattern di percezione categoriale suggerisce che il sistema percettivo non sta semplicemente conducendo un'analisi acustica generale, ma sta applicando categorie fonologiche specifiche del linguaggio.

Il fatto sorprendente è che i neonati, pur non avendo ancora acquisito le categorie fonologiche della loro lingua, dimostrano già percezione categoriale per molti contrasti consonantici. Quando testati con continua sintetici di voice onset time, mostrano pattern di discriminazione simili a quelli degli adulti, con prestazioni elevate per coppie che attraversano il confine categoriale e prestazioni ridotte per coppie omogenee. Questo suggerisce che certi confini categoriali potrebbero non essere appresi dall'esperienza linguistica, ma potrebbero riflettere proprietà intrinseche del sistema percettivo uditivo o addirittura del sistema motorio articolatorio. Una teoria influente, la teoria motoria della percezione del linguaggio proposta da Liberman e colleghi, sostiene che la percezione del linguaggio è fondamentalmente un processo di percezione dei gesti articolatori piuttosto che di analisi acustica pura. Secondo questa prospettiva, quando ascoltiamo il linguaggio, il nostro sistema percettivo recupera automaticamente i comandi motori che avrebbero prodotto quei suoni, e la categorizzazione fonetica riflette la struttura discreta dei gesti articolatori possibili.

Evidenze recenti dalla neuroscienza cognitiva hanno fornito supporto parziale a questa concezione motoria. Studi di neuroimaging funzionale hanno documentato che durante l'ascolto del linguaggio, oltre alle aree uditive primarie nella corteccia temporale superiore, si attivano anche regioni della corteccia premotoria e del giro frontale inferiore, aree tradizionalmente associate con la pianificazione e l'esecuzione di movimenti articolatori. Questa attivazione motoria durante la percezione suggerisce l'esistenza di un sistema di risonanza sensori-motorio che mappa automaticamente i suoni del linguaggio sui gesti articolatori corrispondenti. Studi di stimolazione magnetica transcranica hanno ulteriormente confermato il coinvolgimento causale del sistema motorio nella percezione: quando l'eccitabilità della corteccia motoria viene modulata attraverso stimolazione durante compiti di discriminazione fonetica, si osservano effetti selettivi sulla performance, con facilitazione o interferenza dipendente dalla compatibilità tra i gesti stimolati e i fonemi da discriminare.

Questa architettura sensori-motoria accoppiata potrebbe essere già funzionale, almeno in forma rudimentale, nel neonato. Osservazioni etologiche hanno documentato che neonati di poche settimane sono capaci di imitazione facciale: quando un adulto protrude la lingua o apre la bocca in modo esagerato, il neonato tende a produrre movimenti imitativi, suggerendo l'esistenza di un sistema di mappatura diretta tra percezione visiva di gesti e produzione motoria. Se un meccanismo analogo opera nel dominio uditivo-articolatorio, il neonato potrebbe possedere una predisposizione innata a mappare pattern acustici del linguaggio su schemi motori articolatori prototipici, fornendo una base naturale per la categorizzazione fonetica.

Tuttavia, questa capacità discriminativa universale del neonato non persiste immutata attraverso lo sviluppo. Una scoperta cruciale della ricerca sullo sviluppo percettivo è che intorno ai sei-dodici mesi di vita si verifica un processo di sintonizzazione percettiva: le capacità di discriminazione si affinano per i contrasti rilevanti nella lingua ambientale mentre si deteriorano per i contrasti non nativi. Questo fenomeno è stato documentato in modo

paradigmatico per il contrasto tra le consonanti retroflessse e dentali utilizzate in hindi e in altre lingue indiane, ma assenti nella maggior parte delle lingue europee. Bambini americani di sei-otto mesi discriminano facilmente questo contrasto con prestazioni comparabili a quelle di bambini indiani della stessa età. Tuttavia, entro i dieci-dodici mesi, i bambini americani mostrano un declino drammatico nella capacità di discriminazione, mentre i bambini indiani mantengono o addirittura migliorano le loro prestazioni. Questo pattern suggerisce che l'esperienza percettiva con la lingua ambientale rimodella il sistema fonologico, sintonizzandolo sulle distinzioni rilevanti e riducendo la sensibilità a distinzioni irrilevanti.

Il meccanismo neurale sottostante questa sintonizzazione percettiva può essere concepito come un processo di apprendimento statistico non supervisionato. Il sistema percettivo del bambino è esposto a migliaia di esemplari di sillabe e parole nella lingua ambientale, ciascuno con una realizzazione acustica leggermente diversa a causa della variabilità naturale della produzione vocale. Questa distribuzione di esemplari non è uniforme nello spazio acustico, ma presenta addensamenti e discontinuità che corrispondono alle categorie fonemiche della lingua. Attraverso meccanismi di apprendimento hebbiano operanti nelle reti neurali uditive, le connessioni sinaptiche si rafforzano preferenzialmente per pattern acustici che co-occorrono frequentemente, portando alla formazione di rappresentazioni prototipiche o esemplari-cluster che corrispondono ai fonemi della lingua nativa. Simultaneamente, attraverso processi di competizione tra neuroni e inibizione laterale, il sistema tende a creare confini decisionali netti nello spazio acustico, sharpening le distinzioni tra categorie adiacenti. I contrasti non nativi, non essendo rinforzati dall'esperienza, non sviluppano questa struttura categoriale affinata, e la loro discriminazione si deteriora.

Possiamo formalizzare questo processo utilizzando il framework delle reti neurali auto-organizzanti. Immaginiamo che lo spazio acustico multidimensionale sia rappresentato attraverso una popolazione di neuroni, ciascuno con un campo recettivo specifico, ovvero una regione dello spazio acustico alla quale risponde preferenzialmente. Inizialmente, questi campi recettivi potrebbero essere distribuiti in modo relativamente uniforme, fornendo una copertura ampia dello spazio acustico possibile. Man mano che il bambino viene esposto alla lingua ambientale, la distribuzione degli stimoli acustici modula i pesi sinaptici secondo una regola di apprendimento competitivo. Neuroni che rispondono a regioni dello spazio acustico densamente popolate da esemplari ricevono attivazione frequente e rafforzano le loro connessioni, mentre neuroni sintonizzati su regioni scarsamente popolate ricevono attivazione ridotta e i loro pesi si indeboliscono. Il risultato finale è una mappa auto-organizzata nella quale la densità di rappresentazione neurale riflette la struttura statistica dell'input: regioni dello spazio acustico corrispondenti a fonemi nativi hanno rappresentazioni neurali ricche e differenziate, mentre regioni corrispondenti a fonemi non nativi hanno rappresentazioni impoverite.

Matematicamente, questo può essere descritto attraverso un algoritmo di vector quantization. Sia lo spazio acustico rappresentato come uno spazio vettoriale di dimensione elevata, dove ciascun vettore rappresenta le caratteristiche spettrali di un segmento sonoro. La rete neurale contiene un insieme di vettori prototipo, ciascuno rappresentante il centro di una categoria percettiva. Quando un nuovo stimolo acustico arriva, viene comparato con tutti i prototipi, e il prototipo più vicino viene selezionato come vincitore. Il vettore prototipo vincitore viene quindi aggiornato per avvicinarsi ulteriormente allo stimolo, secondo la regola

di apprendimento: prototipo_nuovo = prototipo_vecchio più eta moltiplicato per la differenza tra stimolo e prototipo_vecchio, dove eta è un tasso di apprendimento piccolo. Questo processo, ripetuto per migliaia di esemplari, porta i prototipi a convergere verso i centri di massa delle distribuzioni di stimoli, effettivamente implementando una forma di clustering non supervisionato. I confini categoriali emergono naturalmente come le regioni di transizione tra bacini di attrazione di prototipi adiacenti.

La Fase del Babbling e la Scoperta della Produzione Vocale

Intorno ai sei-sette mesi di vita, mentre la sintonizzazione percettiva sta procedendo, emerge un nuovo fenomeno cruciale: il babbling canonico. Il bambino inizia a produrre sequenze sillabiche ripetitive come "bababa", "dadada", "mamama", che presentano una struttura temporale notevolmente simile a quella del linguaggio adulto. Queste produzioni non sono più i vocalizzi indifferenziati della fase di cooing precedente, ma rappresentano il primo tentativo sistematico di controllare l'apparato articolatorio per produrre unità discrete simili a sillabe. L'emergere del babbling canonico richiede la maturazione di diversi sistemi neurali che devono coordinarsi in modo preciso: il sistema respiratorio deve fornire un flusso d'aria controllato e modulato, la laringe deve vibrare in modo stabile per produrre fonazione, e gli articolatori sopraglottidici, lingua, labbra, mandibola, devono muoversi rapidamente tra configurazioni diverse per creare le transizioni consonantiche e vocaliche che definiscono la sillaba.

La struttura temporale del babbling canonico è particolarmente significativa. Analisi acustiche dettagliate hanno rivelato che le transizioni tra la fase consonantica e la fase vocalica nelle sillabe balbettate hanno durate dell'ordine di venti-quaranta millisecondi, comparabili alle transizioni nel linguaggio adulto. Questa precisione temporale richiede un controllo motorio sofisticato, poiché i muscoli articolatori devono essere attivati con sequenze temporali precise dell'ordine di decine di millisecondi. La maturazione dei circuiti cerebellari e dei gangli basali, strutture cruciali per il controllo temporale fine del movimento, sembra essere un prerequisito per l'emergere del babbling canonico. Studi su bambini con lesioni cerebellari congenite hanno documentato ritardi significativi nell'onset del babbling e anomalie nella struttura temporale delle produzioni, supportando il ruolo causale di queste strutture nella coordinazione motoria articolatoria.

Crucialmente, il babbling non è semplicemente esercizio motorio casuale, ma mostra fin dall'inizio influenze dell'input linguistico ambientale. Mentre inizialmente il repertorio consonantico del babbling tende a essere simile attraverso lingue diverse, concentrato su consonanti articolatoriamente semplici come le labiali e le dentali-alveolari, entro i nove-dodici mesi emergono differenze sistematiche che riflettono le proprietà fonotattiche e fonologiche della lingua ambientale. Bambini francofoni tendono a produrre più sillabe con vocali anteriori come "i" e "e", riflettendo la maggiore frequenza di queste vocali in francese. Bambini esposti a lingue con ricchi inventari consonantici retroflessi o faringali iniziano a incorporare questi suoni nel loro babbling, mentre bambini esposti a lingue come il giapponese, che ha una struttura sillabica prevalentemente aperta senza cluster consonantici, tendono a evitare sequenze consonantiche nel loro babbling.

Questa convergenza del babbling verso le proprietà della lingua nativa suggerisce che il sistema di produzione vocale sta già operando sotto l'influenza modulatoria del sistema

percettivo. Un modello influente proposto da Patricia Kuhl concepisce questo processo come un loop di feedback uditivo-motorio. Il bambino produce sequenze di babbling e ascolta simultaneamente la propria produzione attraverso conduzione aerea e ossea. Queste rappresentazioni uditive autogenerate vengono comparate con le rappresentazioni percettive dei suoni linguistici ambientali che il bambino ha internalizzato attraverso l'esposizione passiva. Quando si verifica una corrispondenza tra produzione propria e target ambientale, questa associazione sensori-motoria viene rafforzata attraverso meccanismi di apprendimento hebbiano: i neuroni che codificano il pattern motorio articolatorio e i neuroni che codificano il pattern acustico risultante sviluppano connessioni bidirezionali forti. Attraverso ripetizioni successive, il bambino sviluppa gradualmente una mappa di corrispondenze tra configurazioni articolatorie e conseguenze acustiche, una forma rudimentale di quello che diventerà il sistema fonologico produttivo.

Questa concezione del babbling come apprendimento sensori-motorio attraverso feedback uditivo trova supporto drammatico negli studi su bambini sordi. Bambini con sordità profonda congenita iniziano a balbettare approssimativamente alla stessa età dei bambini udenti, suggerendo che l'onset iniziale del babbling è maturazionalmente determinato piuttosto che dipendente dall'input acustico. Tuttavia, il babbling dei bambini sordi mostra differenze qualitative significative: tende a essere meno strutturato ritmicamente, con transizioni temporali meno precise, e non mostra la convergenza verso le proprietà fonetiche della lingua ambientale che caratterizza il babbling dei bambini udenti. Inoltre, mentre il babbling canonico dei bambini udenti tende a diventare progressivamente più abbondante e complesso tra i sette e i dodici mesi, il babbling dei bambini sordi tende a diminuire dopo un picco iniziale, suggerendo che il mantenimento e l'elaborazione del babbling richiedono feedback uditivo. Questi pattern sono consistenti con l'idea che il babbling canonico maturo dipende dall'integrazione sensori-motoria mediata dal feedback uditivo.

Interessantemente, bambini sordi esposti alla lingua dei segni mostrano un fenomeno analogo nel dominio manuale: il babbling manuale. Intorno agli stessi mesi in cui emergono i babbling vocali nei bambini udenti, bambini sordi esposti alla lingua dei segni iniziano a produrre sequenze ritmiche di movimenti manuali che hanno proprietà formali simili ai gesti linguistici della lingua dei segni, ma senza significato comunicativo chiaro. Queste produzioni sono caratterizzate da configurazioni manuali semplici ripetute ritmicamente, distribuite nello spazio segnico canonico utilizzato dai segnanti adulti. Questo parallelismo suggerisce che il babbling riflette un processo generale di esplorazione sensori-motoria nel dominio della comunicazione, indipendente dalla modalità specifica, vocale-uditiva o visivo-gestuale. Il sistema nervoso sembra possedere meccanismi generali per l'apprendimento di mappature sensori-motorie attraverso feedback, che possono essere applicati sia al controllo dell'apparato vocale che al controllo dei movimenti manuali quando questi sono rilevanti per la comunicazione.

L'Emergere della Comprensione Lessicale e la Struttura dell'Input Linguistico

Verso gli otto-dieci mesi, mentre il babbling si sta elaborando, inizia a manifestarsi un fenomeno di portata ancora più profonda: la comprensione di parole. I genitori cominciano a notare che il bambino risponde in modo differenziale a certi suoni linguistici ricorrenti: il nome proprio del bambino elicita orientamento attenzionale, la parola "pappa" in contesti

appropriati suscita eccitazione anticipatoria, la parola "no" pronunciata con intonazione proibitiva produce interruzione comportamentale. Queste osservazioni aneddotiche sono state confermate e quantificate attraverso metodologie sperimentali come il paradigma del looking preference, nel quale il bambino viene posto di fronte a due schermi che mostrano immagini diverse mentre una voce registrata nomina uno degli oggetti raffigurati. Già a nove-dieci mesi, i bambini dimostrano la capacità di orientare lo sguardo preferenzialmente verso l'immagine corrispondente alla parola pronunciata, indicando che hanno stabilito un'associazione tra forma fonologica e referente visivo.

Questa capacità di comprensione lessicale emerge in un contesto comunicativo ricco di supporti contestuali che facilitano l'apprendimento delle associazioni parola-referente. Quando i genitori parlano ai bambini, tendono spontaneamente a utilizzare un registro linguistico modificato, spesso chiamato "motherese" o "infant-directed speech", caratterizzato da proprietà prosodiche e pragmatiche distintive. L'intonazione è esagerata, con contorni pitch più ampi che nel discorso rivolto ad adulti; il tempo di eloquio è rallentato, con pause più lunghe tra enunciati; le parole chiave sono spesso ripetute e enfatizzate attraverso accento prosodico; la referenza è spesso accompagnata da gesti deittici come il pointing che disambiguano l'intenzione comunicativa. Queste modifiche non sono arbitrarie, ma sembrano calibrate in modo ottimale per le capacità percettive e cognitive del bambino. L'intonazione esagerata rende i confini prosodici tra enunciati più salienti, facilitando la segmentazione del flusso continuo in unità discrete. La ripetizione aumenta la frequenza di co-occorrenza tra forme fonologiche e contesti situazionali, rafforzando le associazioni. I gesti disambiguano la referenza, riducendo l'incertezza circa quale aspetto della situazione complessa è rilevante per l'interpretazione della parola.

Tuttavia, anche con questi supporti contestuali, il problema logico dell'induzione lessicale rimane formidabile. Consideriamo una situazione prototipica: un genitore e un bambino osservano insieme un coniglio che salta nel giardino, e il genitore pronuncia l'enunciato "guarda il coniglio". Dal punto di vista del bambino, che non conosce ancora il significato della parola "coniglio", questa situazione è radicalmente sottodeterminata. La parola potrebbe in linea di principio denotare l'animale intero, oppure le sue parti come le orecchie o la coda, oppure il colore del pelo, oppure l'azione di saltare, oppure la configurazione spaziale coniglio-nel-giardino, oppure uno stadio temporale del coniglio, oppure la proprietà astratta di essere-un-coniglio. Quine, nel suo famoso esperimento mentale del linguista radicale che cerca di tradurre una lingua completamente ignota, ha enfatizzato l'indeterminazione radicale di questa situazione: un numero infinito di ipotesi semantiche è logicamente compatibile con qualsiasi insieme finito di osservazioni. Come può allora il bambino convergere sulla interpretazione corretta con l'efficienza e l'affidabilità che osserviamo empiricamente?

La risposta proposta dalla letteratura sullo sviluppo cognitivo è che i bambini non affrontano questo problema come logici puri che considerano tutte le ipotesi logicamente possibili con peso uguale, ma piuttosto applicano vincoli cognitivi innati o apprendimenti precoci che restringono drasticamente lo spazio delle ipotesi considerate, rendendo il problema trattabile. Il vincolo dell'oggetto intero, formulato esplicitamente da Ellen Markman, postula che quando i bambini apprendono un nuovo nome per un'entità, assumono di default che il nome denoti l'oggetto intero piuttosto che parti o proprietà dell'oggetto. Questo vincolo può essere concepito come un bias induttivo che assegna probabilità a priori molto più alte alle

ipotesi che mappano parole su oggetti interi rispetto alle ipotesi che le mappano su parti o attributi. Formalmente, se H_intero rappresenta l'ipotesi che una parola denoti l'oggetto intero e H_parte rappresenta l'ipotesi che denoti una parte, allora il bambino opera con una distribuzione di probabilità a priori tale che P(H_intero) è molto maggiore di P(H_parte), diciamo nell'ordine di cento volte più grande. Conseguentemente, dato un numero limitato di osservazioni della parola in contesti appropriati, l'inferenza bayesiana favorirà fortemente l'ipotesi dell'oggetto intero anche se entrambe le ipotesi sono compatibili con i dati osservati.

Il fondamento cognitivo di questo bias potrebbe risiedere nella natura della segmentazione percettiva precoce. Studi sulla percezione visiva infantile hanno documentato che già nei primi mesi di vita, il sistema visivo segmenta automaticamente le scene in oggetti coesi basandosi su principi gestalitici come la continuità spaziale, il movimento comune, e i confini definiti da discontinuità di texture o colore. Questi processi di segmentazione percettiva producono rappresentazioni di oggetti come unità naturali, e potrebbe essere cognitivamente economico utilizzare queste unità percettive come i referenti di default per nuove parole. Se il lessico è concepito come un sistema di etichette che si mappano su rappresentazioni concettuali, e se le rappresentazioni concettuali più salienti e accessibili sono quelle di oggetti interi piuttosto che di parti o attributi, allora l'assunzione dell'oggetto intero emerge naturalmente come strategia di apprendimento di default.

Un secondo vincolo cruciale è l'assunzione tassonomica, che guida i bambini a estendere nuove parole a categorie tassonomiche basate su similarità di tipo piuttosto che a raggruppamenti tematici basati su associazioni situazionali. Quando un bambino apprende che un particolare esemplare è chiamato "cane", deve inferire quali altri oggetti nel mondo possono essere appropriatamente etichettati con quella parola. Una possibilità sarebbe estendere la parola a oggetti che co-occorrono frequentemente con cani in situazioni tipiche: ciotole per cibo, guinzagli, cucce. Questa sarebbe un'estensione tematica. Alternativamente, il bambino potrebbe estendere la parola a oggetti che condividono proprietà percettive e funzionali fondamentali con il cane originale: altri animali quadrupedi con pelo, coda, che abbaiano. Questa sarebbe un'estensione tassonomica. Le evidenze empiriche dimostrano chiaramente che i bambini preferiscono l'estensione tassonomica. Quando viene insegnata una nuova parola per un oggetto, e poi viene chiesto al bambino di identificare altri esemplari della stessa categoria, i bambini selezionano oggetti tassonomicamente simili piuttosto che oggetti tematicamente associati, anche quando entrambi i tipi di oggetti sono ugualmente familiari.

Questa preferenza tassonomica non sembra essere semplicemente un artefatto delle procedure sperimentali, poiché emerge anche in compiti più naturalistici. Quando i bambini iniziano a produrre parole spontaneamente, le loro sovraestensioni tendono a essere tassonomiche: "cane" viene erroneamente applicato a gatti, mucche, cavalli, altri animali quadrupedi, ma raramente a cucce o ciotole. Similmente, "palla" viene sovraesteso a oggetti sferici di vario tipo, ma non tipicamente a oggetti associati tematicamente al gioco con la palla. Questo pattern suggerisce che il bias tassonomico riflette una struttura profonda del sistema concettuale piuttosto che una strategia superficiale dipendente dal compito.

Dal punto di vista computazionale, il bias tassonomico può essere interpretato come la preferenza per ipotesi che massimizzano la coesione intra-categoria e la separazione inter-categoria nello spazio di similarità percettiva. Le categorie tassonomiche tendono a

corrispondere a regioni dense e compatte nello spazio multidimensionale delle features percettive: tutti i cani condividono un'ampia sovrapposizione di features come la presenza di pelo, quattro zampe, musello, coda, capacità di movimento autonomo, vocalizzi caratteristici. Al contrario, i raggruppamenti tematici tendono a essere sparsi e eterogenei nello spazio percettivo: una cuccia, un guinzaglio e una ciotola condividono pochissime features percettive dirette, essendo connessi soltanto dalla loro relazione funzionale con i cani. Se il sistema di apprendimento lessicale è configurato per preferire ipotesi che corrispondono a regioni compatte nello spazio di similarità, emergerà naturalmente una preferenza per estensioni tassonomiche.

Un terzo vincolo che facilita potentemente l'apprendimento lessicale è il principio di mutua esclusività, secondo cui ogni oggetto dovrebbe avere un solo nome al livello categoriale di base. Questo principio genera inferenze potenti in situazioni di ambiguità referenziale. Consideriamo il fenomeno del fast mapping, documentato in modo paradigmatico da Susan Carey e Elsa Bartlett. Un bambino che conosce già la parola "tazza" viene posto in una situazione dove due oggetti sono visibili: una tazza familiare e un oggetto nuovo mai visto prima, per esempio una pinza da cucina. L'adulto pronuncia una richiesta contenente una parola inventata: "per favore, dammi il dax". Notevolmente, bambini già a due anni di età selezionano sistematicamente l'oggetto nuovo, la pinza, inferendo che la parola sconosciuta deve riferirsi all'oggetto per il quale non hanno ancora un nome. Il ragionamento inferenziale sembra procedere attraverso un processo di eliminazione: la tazza ha già un nome, "tazza"; la nuova parola "dax" deve quindi riferirsi a qualcos'altro; l'unico altro oggetto saliente nella situazione è la pinza; quindi "dax" probabilmente denota la pinza.

Questo processo può essere formalizzato come inferenza pragmatica guidata dal principio di cooperazione comunicativa. Se assumiamo che il parlante è cooperativo e sta cercando di comunicare informazione rilevante e informativa, l'uso di una parola nuova implica l'intenzione di riferirsi a qualcosa per cui il bambino non ha già un'etichetta lessicale. Altrimenti, il parlante avrebbe usato la parola conosciuta. Formalmente, possiamo modellare questo attraverso il calcolo delle implicature scalari in pragmatica griciana. Sia "tazza" un'alternativa lessicale più informativa di "dax" per riferirsi alla tazza, nel senso che il bambino conosce già "tazza". Allora, se il parlante dice "dax" invece di "tazza", il bambino inferisce che il parlante intendeva comunicare qualcosa che non può essere comunicato usando "tazza", ovvero che il referente non è una tazza. Questa inferenza elimina la tazza come possibile referente, lasciando la pinza come unico candidato.

Il principio di mutua esclusività non deve essere interpretato come un vincolo logico assoluto che proibisce sinonimia o gerarchia tassonomica, poiché ovviamente le lingue contengono molteplici parole che si applicano allo stesso oggetto a diversi livelli di specificità: "animale", "cane", "golden retriever" possono tutti riferirsi alla stessa entità. Piuttosto, dovrebbe essere concepito come una presunzione di default che opera principalmente al livello categoriale di base e che può essere superata da evidenza contraria. Quando un bambino inizialmente apprende "cane", il principio di mutua esclusività impedisce l'inferenza che "animale" o "golden retriever" si riferiscono esattamente alla stessa categoria, facilitando la differenziazione lessicale. Successivamente, attraverso esposizione a usi linguistici che violano la mutua esclusività, come "questo cane è un golden retriever" o "i cani sono animali", il bambino può apprendere le relazioni gerarchiche tassonomiche e accettare la co-applicabilità di molteplici termini.

La Dinamica Oscillatoria come Substrato Computazionale del Linguaggio

L'Emergere di un Paradigma: Dalle Rappresentazioni Statiche alle Dinamiche Temporali

Per decenni, la neuroscienza cognitiva del linguaggio si è concentrata prevalentemente sull'identificazione di regioni cerebrali specializzate per diverse funzioni linguistiche, concependo il cervello come una collezione di moduli anatomicamente distinti ciascuno dedicato a un aspetto specifico dell'elaborazione linguistica. L'area di Broca per la sintassi e la produzione, l'area di Wernicke per la comprensione semantica, il giro angolare per l'elaborazione delle dipendenze a lunga distanza: questa concezione modulare, pur avendo fornito intuizioni preziose e corrispondenze approssimative con pattern lesionali, lascia irrisolto un problema fondamentale. Il linguaggio è intrinsecamente un fenomeno temporale: le parole e le strutture sintattiche si dispiegano nel tempo con una organizzazione gerarchica precisa, e la comprensione richiede il mantenimento simultaneo di molteplici rappresentazioni parziali che devono essere integrate e aggiornate dinamicamente man mano che l'input procede. Come possono popolazioni di neuroni distribuiti attraverso molteplici regioni corticali coordinare la loro attività in modo da implementare questa complessa coreografia computazionale?

La risposta che emerge dalla ricerca contemporanea in elettrofisiologia cognitiva è che le oscillazioni neurali, ovvero le fluttuazioni ritmiche dell'eccitabilità e dell'attività sincronizzata di popolazioni neuronali, costituiscono un meccanismo fondamentale attraverso cui il cervello organizza l'elaborazione dell'informazione nel tempo e coordina la comunicazione tra regioni distribuite. Queste oscillazioni non sono semplicemente epifenomeni dell'attività neurale, ma rappresentano una caratteristica computazionalmente rilevante dell'architettura neurale che permette l'implementazione di operazioni come il binding temporale, la segmentazione di sequenze, il multiplexing di rappresentazioni multiple, e la modulazione dinamica della connettività funzionale.

Le oscillazioni neurali sono misurabili attraverso tecniche di registrazione dell'attività elettrica cerebrale come l'elettroencefalografia e la magnetoencefalografia, che catturano con risoluzione temporale millimetrica le fluttuazioni nei campi elettrici e magnetici generati dall'attività sincronizzata di grandi popolazioni neuronali. Quando migliaia di neuroni corticali organizzati in architetture colonnari scaricano in modo sincronizzato, i loro campi elettrici locali si sommano generando potenziali misurabili a distanza, sul cuoio capelluto o attraverso sensori magnetici. L'analisi di questi segnali nel dominio della frequenza, tipicamente attraverso trasformata di Fourier o wavelet, rivela che l'attività neurale contiene componenti oscillatorie robuste in diverse bande di frequenza, ciascuna associata con stati cognitivi e computazioni specifiche.

Le Oscillazioni Delta e l'Aggancio alla Struttura Prosodica del Discorso

Le oscillazioni nella banda delta, caratterizzate da frequenze nell'intervallo da uno a quattro hertz, corrispondono a periodicità nell'ordine di duecentocinquanta millisecondi fino a un secondo, scale temporali che coincidono notevolmente con la struttura temporale delle frasi

e degli enunciati nel linguaggio parlato. Una frase tipica può durare da uno a tre secondi, con confini frasali marcati prosodicamente da variazioni sistematiche di pitch, intensità e durata delle sillabe finali. Questa corrispondenza temporale non è casuale: una mole crescente di evidenze suggerisce che le oscillazioni delta nella corteccia uditiva si agganciano, ovvero si sincronizzano, con la struttura prosodica del discorso in arrivo, con i picchi dell'oscillazione delta tendenti a coincidere con i confini tra enunciati e le valli corrispondenti alle porzioni centrali degli enunciati.

Il fenomeno dell'entrainment, ovvero dell'aggancio di fase tra oscillazioni neurali e ritmi esterni, è stato documentato attraverso misure di coerenza tra l'envelope di ampiezza del segnale acustico del discorso e l'attività oscillatoria delta registrata dalla corteccia temporale superiore. Quando i partecipanti ascoltano discorso continuo, si osserva una coerenza significativa tra le due serie temporali: la fase dell'oscillazione delta predice in modo statisticamente affidabile la presenza di confini prosodici nel segnale acustico. Questa relazione non è meramente epifenomenale: esperimenti di manipolazione causale attraverso stimolazione transcranica con corrente alternata, nella quale correnti elettriche deboli vengono applicate al cuoio capelluto per modulare ritmicamente l'eccitabilità corticale, hanno dimostrato che modulare artificialmente la fase delle oscillazioni delta può facilitare o interferire con la comprensione del discorso in modo dipendente dalla fase. Quando la stimolazione delta è sincronizzata appropriatamente con la struttura prosodica del discorso, si osserva facilitazione della comprensione; quando è asincrona o in controfase, si osserva interferenza.

Ma quale funzione computazionale serve questo aggancio delta alla prosodia? Una possibilità teoricamente plausibile è che le oscillazioni delta implementino un meccanismo di predizione temporale, permettendo al sistema di anticipare quando informazione critica arriverà e di allocare risorse attenzionali di conseguenza. La struttura prosodica del linguaggio è altamente regolare e predicibile: i confini frasali sono marcati da allungamenti temporali, pause, e pattern intonativi stereotipati. Se il sistema percettivo può tracciare questa struttura ritmmica attraverso oscillazioni delta aggancianti, la fase corrente dell'oscillazione fornisce informazione circa la posizione temporale all'interno della frase, permettendo predizioni circa l'imminenza di confini strutturali. Quando l'oscillazione delta si avvicina al suo picco, il sistema "sa" implicitamente che un confine frasale è probabilmente imminente e può prepararsi per operazioni di integrazione semantica e reset di buffer sintattici.

Possiamo formalizzare questa concezione attraverso un modello di oscillatore di fase accoppiato all'input ritmico. Rappresentiamo la fase dell'oscillazione delta come una variabile angolare phi che evolve nel tempo secondo un'equazione differenziale: la derivata temporale di phi uguale a omega più K moltiplicato per il seno della differenza tra phi_input e phi, dove omega è la frequenza naturale intrinseca dell'oscillatore, phi_input è la fase estratta dall'envelope del segnale acustico, e K è la forza di accoppiamento. Questa equazione di Kuramoto descrive come l'oscillatore neurale si aggancia all'input: quando la fase neurale è in ritardo rispetto alla fase dell'input, il termine di accoppiamento accelera l'oscillazione; quando è in anticipo, la rallenta. Il risultato è una convergenza verso uno stato di sincronizzazione stabile nel quale phi rimane costantemente in una relazione di fase fissa con phi_input.

L'utilità computazionale di questa sincronizzazione diventa evidente quando consideriamo che la fase dell'oscillazione modula ciclicamente l'eccitabilità dei neuroni. Durante la fase ascendente verso il picco, i neuroni sono iperpolarizzati e meno responsivi agli input; durante la fase discendente verso il minimo, sono depolarizzati e massimamente eccitabili. Di conseguenza, input che arrivano in fasi diverse dell'oscillazione riceveranno gradi diversi di elaborazione: input che coincidono con fasi di alta eccitabilità verranno processati vigorosamente, mentre input che cadono in fasi di bassa eccitabilità verranno attenuati. Se l'oscillazione è sincronizzata con la struttura prosodica tale che i picchi corrispondono ai confini frasali, allora l'eccitabilità sarà massima precisamente quando arrivano i marcatori prosodici di confine, facilitando la loro rilevazione e l'innescamento di processi di integrazione.

Le Oscillazioni Theta come Organizzatori di Chunking Lessicale e Sintattico

Salendo nella gerarchia di scale temporali, le oscillazioni theta, con frequenze da quattro a otto hertz corrispondenti a periodicità da centocinquanta a duecentocinquanta millisecondi, operano a una scala temporale che corrisponde approssimativamente alla durata di parole individuali o sintagmi brevi. Un ciclo theta di duecento millisecondi può contenere una parola bisillabica o trisillabica pronunciata a velocità normale, suggerendo che theta potrebbe organizzare l'elaborazione a livello di unità lessicali o sintattiche elementari. Le evidenze empiriche supportano fortemente questa ipotesi: la potenza delle oscillazioni theta nella corteccia frontale e temporale aumenta sistematicamente durante l'elaborazione di frasi sintatticamente complesse rispetto a sequenze di parole semplici o non strutturate. Questo incremento di potenza theta non riflette semplicemente un aumento generale dell'effort cognitivo, poiché è selettivo per complessità sintattica piuttosto che per altre forme di difficoltà come la frequenza lessicale bassa o la plausibilità semantica ridotta.

Un aspetto particolarmente intrigante delle oscillazioni theta è il loro ruolo nel supportare la memoria di lavoro, la capacità di mantenere attivamente informazione per brevi periodi durante l'elaborazione cognitiva. Gli studi classici di George Miller negli anni cinquanta del Novecento documentarono che la capacità della memoria di lavoro umana è limitata a circa sette più o meno due items, un numero che è rimasto misterioso per decenni. Perché proprio questo numero magico? Una spiegazione affascinante emerge dalla considerazione delle oscillazioni theta. Se ogni ciclo theta può codificare o mantenere un chunk di informazione attraverso l'attivazione sincronizzata di una assembly neuronale specifica durante quel ciclo, allora la capacità totale della memoria di lavoro dipende da quanti cicli theta possono essere "impilati" nella finestra temporale durante la quale l'informazione deve essere mantenuta.

Formalizziamo questa idea. Supponiamo che l'informazione debba essere mantenuta per una durata T, diciamo circa due secondi, che è una stima ragionevole della finestra temporale per la memoria di lavoro verbale. Se la frequenza theta è f_theta, circa cinque hertz, allora il numero di cicli theta che si verificano durante T è N uguale a f_theta moltiplicato per T, che risulta cinque moltiplicato per due uguale a dieci cicli. Se ogni ciclo può codificare un item, otteniamo una capacità di circa dieci items. Tuttavia, vincoli addizionali riducono questa stima. La discriminabilità tra items richiede che gli stati neurali corrispondenti a items diversi siano sufficientemente distinti. Se items vengono codificati

attraverso pattern di attivazione sincronizzata in fasi diverse del ciclo theta, allora la risoluzione di fase limita il numero di items distinguibili. Con una risoluzione di fase di circa quarantacinque gradi, otteniamo otto stati distinguibili per ciclo. Ma complessivamente, attraverso l'intera finestra temporale, possono verificarsi interferenze tra items che occupano cicli diversi, riducendo ulteriormente la capacità effettiva a circa sette più o meno due items, convergendo notevolmente con la stima empirica di Miller.

Questa concezione della memoria di lavoro come multiplexing temporale attraverso cicli theta genera predizioni testabili. Una predizione cruciale è che la capacità della memoria di lavoro dovrebbe essere modulabile attraverso manipolazione della frequenza theta. Se la frequenza theta viene aumentata, per esempio attraverso stimolazione elettrica ritmica, ciascun ciclo diventa più breve ma più cicli si verificano nella finestra temporale fissa, potenzialmente aumentando la capacità. Viceversa, ridurre la frequenza theta dovrebbe ridurre la capacità. Esperimenti preliminari utilizzando stimolazione transcranica con corrente alternata a frequenze theta hanno fornito supporto parziale a questa predizione, con modulazioni selettive della performance in task di memoria di lavoro dipendenti dalla frequenza di stimolazione.

La Composizione delle Rappresentazioni Multi-Componenziali

Avendo stabilito che le rappresentazioni semantiche lessicali sono strutture multi-componenziali che includono non soltanto informazione simbolica astratta ma anche informazione percettiva, motoria, affettiva ed episodica, dobbiamo ora affrontare la questione cruciale di come queste rappresentazioni complesse si compongono quando combiniamo parole in sintagmi e frasi. La semantica compositiva tradizionale fornisce regole precise per comporre le componenti simboliche attraverso operazioni come l'applicazione funzionale, ma non dice nulla riguardo a come le componenti grounded dovrebbero combinarsi.

Consideriamo un sintagma aggettivo-nome semplice come "mela rossa". La semantica formale standard ci dice che la componente simbolica del sintagma è ottenuta attraverso la modificazione predicativa: se la componente simbolica di "mela" è S_mela uguale a λx punto quadrata Mela di x chiusa quadra di tipo ⟨e,t⟩ e la componente simbolica di "rossa" è S_rossa uguale a λx punto quadrata Rosso di x chiusa quadra anche di tipo ⟨e,t⟩, allora la componente simbolica del sintagma combinato è S_mela_rossa uguale a λx punto quadrata Mela di x e Rosso di x chiusa quadra, che denota l'insieme delle entità che sono simultaneamente mele e rosse. Questa operazione di intersezione predicativa è ben compresa e formalizzata nella semantica di tipo.

Ma cosa succede alle componenti percettive? Come si combina il vettore percettivo P_mela che codifica features visive, tattili, gustative e olfattive tipiche delle mele con il contributo percettivo di "rossa" per produrre una rappresentazione percettiva del sintagma completo? Una possibilità è che la componente percettiva dell'aggettivo funziona come un operatore di modulazione che modifica selettivamente certe dimensioni del vettore percettivo del nome. Formalmente, possiamo modellare questo processo come un'operazione di addizione vettoriale pesata: P_mela_rossa uguale a P_mela più α moltiplicato per ΔP_colore_rosso, dove P_mela è il vettore percettivo base per il concetto MELA, ΔP_colore_rosso è un vettore

di modificazione che ha componenti non-zero soltanto sulle dimensioni che corrispondono a proprietà di colore e che specifica come queste dimensioni dovrebbero essere modificate per riflettere rossezza, e α è un parametro scalare che determina la forza della modulazione, tipicamente nell'intervallo da zero a uno.

Più concretamente, il vettore P_mela potrebbe avere una distribuzione di valori sulle dimensioni di colore che riflette la variabilità naturale nel colore delle mele, con pesi moderati su rosso, verde e giallo. Il vettore di modulazione ΔP_colore_rosso avrebbe un valore positivo elevato sulla dimensione che codifica rossezza, valori negativi sulle dimensioni che codificano verdezza e giallezza, e valori zero su tutte le altre dimensioni. Quando applichiamo l'operazione di modulazione, le dimensioni di colore nel vettore risultante vengono spostate verso maggiore rossezza e ridotta verdezza e giallezza, producendo una rappresentazione percettiva che corrisponde più strettamente alle mele rosse specificamente piuttosto che alle mele in generale.

Formalmente, se P_mela ha componenti P_mela aperta quadra i chiusa quadra per ciascuna dimensione i, e ΔP_colore_rosso ha componenti ΔP_colore_rosso aperta quadra i chiusa quadra, allora la componente i-esima del vettore risultante è: P_mela_rossa aperta quadra i chiusa quadra uguale a P_mela aperta quadra i chiusa quadra più α moltiplicato per ΔP_colore_rosso aperta quadra i chiusa quadra. Per dimensioni di colore dove i appartiene all'insieme delle dimensioni di colore, il vettore delta ha componenti non-zero che implementano la modificazione: ΔP_colore_rosso aperta quadra rosso chiusa quadra uguale a più uno virgola zero, ΔP_colore_rosso aperta quadra verde chiusa quadra uguale a meno zero virgola cinque, ΔP_colore_rosso aperta quadra giallo chiusa quadra uguale a meno zero virgola tre. Per tutte le altre dimensioni dove i non appartiene alle dimensioni di colore, abbiamo ΔP_colore_rosso aperta quadra i chiusa quadra uguale a zero, lasciando quelle dimensioni invariate.

Dopo l'applicazione di questa operazione di modulazione con un valore appropriato di α, per esempio α uguale a zero virgola sette, otteniamo un vettore percettivo P_mela_rossa che ha feature visive molto simili a P_mela per quanto riguarda forma, texture, dimensione e altre proprietà non cromatiche, ma con una distribuzione di colore spostata decisamente verso il rosso. Questo vettore corrisponde quindi a una simulazione mentale visiva di una mela rossa specifica.

Un principio compositivo analogo può essere applicato per la combinazione di componenti motorie in costruzioni verbo-oggetto. Consideriamo la frase "lanciare la palla". La componente motoria V_lanciare specifica gli schemi motori associati all'azione di lanciare, che includono sequenze come preparazione_braccio, accelerazione_braccio, rilascio_oggetto, e follow-through. La componente motoria V_palla specifica le affordances della palla, che includono schemi come afferrare_con_configurazione_sferica, rimbalzare, rotolare. La combinazione di queste due componenti attraverso la relazione verbo-oggetto produce una rappresentazione motoria integrata che specifica l'intera sequenza motoria di lanciare una palla specifica.

Formalmente, possiamo modellare questa operazione di unificazione come: V_lanciare_palla uguale a Unifica aperta tonda V_lanciare, V_palla chiusa tonda, dove Unifica è un operatore che seleziona schemi motori compatibili da entrambi gli insiemi e li

sequenza appropriatamente. L'algoritmo di unificazione potrebbe procedere come segue: primo, dagli schemi in V_palla, seleziona lo schema di afferramento che è compatibile con l'oggetto sferico, ovvero afferrare_con_configurazione_sferica. Secondo, dagli schemi in V_lanciare, seleziona la sequenza motoria di lanciare che inizia con la configurazione di presa. Terzo, concatena questi schemi nella sequenza temporale appropriata: afferrare_palla_sferica seguito da preparazione_braccio seguito da accelerazione_braccio seguito da rilascio seguito da follow-through seguito dalla traiettoria_parabolica della palla lanciata.

Questa sequenza motoria unificata costituisce essenzialmente una simulazione mentale dell'azione completa di lanciare la palla, una simulazione che include non soltanto i movimenti del corpo ma anche la traiettoria prevista dell'oggetto dopo il rilascio. Questa simulazione può essere utilizzata per predire le conseguenze dell'azione, per pianificare l'esecuzione motoria se l'azione deve essere effettivamente eseguita, o semplicemente per comprendere il significato della frase in termini di un modello ricco e situato dell'evento descritto.

La Costruzione di Simulazioni Dinamiche per Frasi Evento

Per frasi intere che descrivono eventi complessi con molteplici partecipanti e relazioni dinamiche, la rappresentazione semantica grounded corrisponde a una simulazione mentale dinamica che si svolge nel tempo cognitivo, rappresentando l'evoluzione dell'evento descritto. Consideriamo la frase "Il gatto insegue il topo", che descrive un evento di inseguimento predatorio con una struttura temporale estesa e una dinamica spaziale complessa.

La costruzione della simulazione mentale per questa frase procede attraverso molteplici componenti integrate. La componente visuale della simulazione include rappresentazioni visive dinamiche dei due agenti, il gatto e il topo, con le loro features percettive caratteristiche come dimensione, forma, texture del pelo, colorazione. Queste rappresentazioni visive non sono statiche ma cambiano nel tempo simulato per riflettere il movimento relativo degli agenti. Il topo è rappresentato come fuggente con una traiettoria spaziale che si allontana progressivamente da una locazione di partenza, con movimenti rapidi e erratici che sono caratteristici dei comportamenti di fuga dei roditori. Il gatto è rappresentato come inseguitore con una traiettoria che segue e tenta di intercettare la traiettoria del topo, con movimenti che riflettono i pattern motori felini di caccia come scatti rapidi, salti e cambi improvvisi di direzione.

La rappresentazione include anche features cinematiche quantitative come velocità istantanea e accelerazione per ciascun agente. Tipicamente, nella simulazione di un inseguimento, il predatore ha velocità massima superiore ma accelerazione iniziale inferiore rispetto alla preda, creando dinamiche interessanti dove la preda può sfuggire attraverso cambi rapidi di direzione che sfruttano la sua maggiore agilità anche se il predatore è più veloce in linea retta. Queste features cinematiche possono essere rappresentate formalmente come funzioni del tempo: posizione_gatto di t uguale a vettore x_g di t, y_g di t,

z_g di t chiuso vettore e posizione_topo di t uguale a vettore x_t di t, y_t di t, z_t di t chiuso vettore, con le derivate temporali che forniscono velocità e accelerazione.

La componente motoria della simulazione è più complessa perché dipende dalla prospettiva adottata nella simulazione. Se il simulatore adotta la prospettiva del gatto, allora la componente motoria include i pattern motori di inseguimento felino, con rappresentazioni di tensione muscolare nella preparazione al salto, le forze applicate durante lo scatto, la coordinazione visuo-motoria richiesta per tracciare la preda in movimento e predire la sua traiettoria futura. Se invece il simulatore adotta la prospettiva del topo, la componente motoria include i pattern motori di fuga con elevata frequenza di cambi direzionali, accelerazioni massime sfruttando la muscolatura posteriore potente, e continua vigilanza visiva per monitorare la posizione del predatore.

Evidenze sperimentali suggeriscono che i lettori possono adottare fluidamente diverse prospettive durante la comprensione di narrativa, e che la prospettiva adottata influenza quali componenti motorie vengono attivate. Studi che hanno utilizzato paradigmi di compatibilità motoria hanno dimostrato che dopo leggere frasi di azione descritte dalla prospettiva di prima persona, i partecipanti mostrano facilitazione per movimenti compatibili con l'esecuzione dell'azione descritta, mentre dopo leggere le stesse azioni descritte dalla prospettiva di terza persona esterna, l'effetto è ridotto o assente, suggerendo che la simulazione motoria è modulata dalla prospettiva narrativa.

La componente affettiva della simulazione per "Il gatto insegue il topo" include stati emotivi distinti per i due partecipanti. Dal punto di vista del topo, la simulazione include stati affettivi di paura intensa, ansia, arousal elevato associato alla risposta di lotta o fuga del sistema nervoso simpatico. Dal punto di vista del gatto, include stati di eccitazione predatoria, focus attenzionale intenso sulla preda, arousal elevato ma di carattere diverso rispetto alla paura della preda, più simile all'eccitazione anticipatoria associata all'attività di caccia.

Formalmente, possiamo rappresentare la simulazione completa come una traiettoria nello spazio di stati multi-dimensionale che include tutte queste componenti che variano nel tempo: Simulazione di t uguale a aperta graffa P_visuale di t, V_motoria di t, A_affettiva di t chiusa graffa per t appartenente all'intervallo aperta quadra zero, T chiusa quadra, dove t varia da zero che rappresenta l'inizio dell'evento simulato a T che rappresenta la conclusione dell'evento simulato, e la durata T del tempo simulato può non corrispondere linearmente al tempo fisico di processamento. Le evidenze suggeriscono che le simulazioni mentali possono procedere a velocità variabili, talvolta accelerate rispetto al tempo reale per eventi estesi, talvolta rallentate o addirittura fermate per permettere elaborazione dettagliata di momenti critici.

L'Architettura Duale: Sistemi Verbale-Simbolico e Immaginativo-Percettivo

Una teoria unificata della comprensione linguistica che integra coerentemente le evidenze per rappresentazioni simboliche astratte con le evidenze per rappresentazioni grounded percettive e motorie richiede un'architettura cognitiva che accomodi entrambi i tipi di rappresentazione e specifica le loro relazioni funzionali. Un framework teorico influente che

fornisce precisamente questa architettura è la teoria del dual-coding sviluppata originalmente da Allan Paivio negli anni settanta e successivamente elaborata e raffinata attraverso decenni di ricerca empirica.

Secondo la teoria del dual-coding nella sua formulazione contemporanea, il sistema cognitivo umano mantiene due sistemi rappresentazionali parzialmente indipendenti ma fortemente interconnessi che operano in parallelo durante la comprensione linguistica. Il primo sistema, che possiamo chiamare il sistema verbale-simbolico, elabora informazione in un formato proposizionale astratto che è essenzialmente amodale, ovvero non preserva le proprietà percettive specifiche delle esperienze sensoriali originali. Questo sistema implementa rappresentazioni simboliche del tipo che la semantica formale tradizionale postula, con strutture proposizionali composte da predicati, argomenti, quantificatori e operatori logici. Il sistema verbale-simbolico supporta ragionamento inferenziale formale, permettendo la derivazione di conclusioni logiche attraverso regole di inferenza deduttiva, e è implementato neuralmente in reti corticali che includono particolarmente le regioni del giro frontale inferiore sinistro, del giro temporale medio e superiore sinistro, e delle regioni del polo temporale anteriore che sono state associate con elaborazione semantica astratta.

Il secondo sistema, che possiamo chiamare il sistema immaginativo-percettivo, elabora informazione in formati modali multipli che preservano la topologia e le proprietà analogiche delle esperienze percettive e motorie originali. Questo sistema include rappresentazioni visive che mantengono relazioni spaziali e proprietà visive, rappresentazioni auditive che preservano struttura temporale e features acustiche, rappresentazioni somatosensoriali e motorie che codificano configurazioni corporee e sequenze di azioni, e rappresentazioni in altre modalità. Il formato di rappresentazione in questo sistema è analogico nel senso che le relazioni strutturali nelle rappresentazioni corrispondono isomorficamente a relazioni strutturali nei domini rappresentati. Per esempio, distanze spaziali relative tra oggetti in una scena visiva immaginata corrispondono a distanze relative nelle rappresentazioni neurali attraverso la corteccia visiva, preservando quindi la metrica spaziale. Questo sistema supporta simulazioni mentali, immaginazione visiva, e ragionamento spaziale, ed è implementato neuralmente in regioni della corteccia sensoriale e motoria che sono organizzate topograficamente, includendo la corteccia visiva primaria e extrastriata, la corteccia uditiva, la corteccia somatosensoriale e la corteccia motoria.

I due sistemi sono interconnessi attraverso connessioni referenziali bidirezionali che permettono l'attivazione di rappresentazioni in un sistema di innescare l'attivazione di rappresentazioni corrispondenti nell'altro sistema. Quando una parola viene letta o ascoltata, essa attiva inizialmente una rappresentazione lessicale nel sistema verbale che include informazione fonologica, ortografica e sintattica. Questa rappresentazione lessicale è connessa attraverso links referenziali a una rappresentazione semantica verbale che specifica il significato della parola in formato proposizionale astratto. Simultaneamente o con latenza breve, attraverso links referenziali addizionali, la rappresentazione lessicale attiva anche rappresentazioni percettive corrispondenti nel sistema immaginativo. Per una parola concreta come "tazza", questi links attiverebbero rappresentazioni visive della forma tipica di una tazza, rappresentazioni motorie delle azioni di afferrare e bere, e altre rappresentazioni modali associate.

Formalmente, possiamo modellare questa architettura duale specificando due spazi rappresentazionali distinti con morfismi che li connettono. Sia R_verbale lo spazio delle rappresentazioni verbali-simboliche, e sia R_immaginativo lo spazio delle rappresentazioni immaginative-percettive. Per ciascun concetto lessicale C, esistono due rappresentazioni: r_verbale di C appartenente a R_verbale e r_immaginativo di C appartenente a R_immaginativo. Questi sono connessi attraverso una funzione di mappatura bidirezionale: φ: R_verbale ↔ R_immaginativo tale che φ di r_verbale di C uguale a r_immaginativo di C e φ inversa di r_immaginativo di C uguale a r_verbale di C.

L'attivazione di un concetto C nel sistema cognitivo produce attivazione in entrambi i sistemi, con il livello totale di attivazione che è una combinazione pesata delle attivazioni nei due sistemi: Attivazione_totale di C uguale a Attivazione_verbale di C più β moltiplicato per Attivazione_immaginativa di C, dove β è un parametro di ponderazione che modula quanto fortemente le rappresentazioni immaginative contribuiscono all'attivazione totale. Il valore di β non è fisso ma varia dinamicamente con il task cognitivo corrente e con il contesto. Durante task che richiedono ragionamento logico astratto o elaborazione sintattica complessa, β tende a essere relativamente basso, permettendo al sistema verbale-simbolico di dominare. Durante task che richiedono immaginazione visiva dettagliata, navigazione spaziale o simulazione di azioni motorie, β tende a essere elevato, permettendo al sistema immaginativo di contribuire sostanzialmente.

Evidenze per questa modulazione dinamica del peso relativo dei due sistemi provengono da studi di neuroimaging che hanno manipolato sistematicamente la concretezza o l'imageability del materiale linguistico processato. Quando i partecipanti leggono o ascoltano frasi altamente concrete con alto imageability, come "Il bambino ha lanciato la palla rossa nel giardino", gli studi fMRI rivelano attivazione robusta non soltanto nelle regioni linguistiche classiche del sistema verbale ma anche in regioni della corteccia visiva che elaborano colore e movimento, in regioni della corteccia motoria associate con movimenti di lancio del braccio, e in regioni del lobo parietale associate con rappresentazioni spaziali. Quando invece i partecipanti leggono frasi altamente astratte con basso imageability, come "La verità è relativa alla prospettiva epistemica", l'attivazione è predominante nelle regioni del sistema verbale con attivazione minima o assente nelle cortecce sensoriali e motorie.

Questa evidenza supporta l'idea che il parametro β è modulato dalla concretezza e imageability del materiale: β alto per materiale concreto, β basso per materiale astratto. Formalmente, possiamo modellare questa dipendenza come: β uguale a f aperta tonda Concretezza, Imageability, Task_Demands chiusa tonda, dove f è una funzione crescente della concretezza e imageability del materiale linguistico e modulata dalle richieste del task cognitivo corrente.

Le Predizioni Testabili del Modello Integrato

Un modello teorico scientifico deve generare predizioni empiriche specifiche e testabili che potrebbero potenzialmente falsificare il modello se non fossero confermate. Il modello integrato multi-componenziale della semantica che abbiamo sviluppato genera numerose predizioni precise riguardo a pattern di dissociazione neurologica, effetti comportamentali e pattern di attivazione neurale.

La prima predizione centrale riguarda pattern di doppia dissociazione tra elaborazione simbolica e elaborazione percettiva in pazienti neurologici con lesioni cerebrali focali. Se i sistemi verbale-simbolico e immaginativo-percettivo sono parzialmente indipendenti e implementati in regioni neurali distinte, allora dovremmo osservare pazienti con lesioni selettive a uno dei due sistemi che mostrano deficit corrispondentemente selettivi. Specificamente, pazienti con lesioni estese nelle regioni linguistiche classiche come il giro frontale inferiore sinistro e le regioni temporali sinistre dovrebbero mostrare deficit sostanziali in task che richiedono elaborazione simbolica verbale, come giudizi di relazioni tassonomiche semantiche basate su definizioni verbali, comprensione di frasi sintatticamente complesse, o elaborazione di linguaggio astratto. Tuttavia, questi stessi pazienti dovrebbero mostrare performance relativamente preservata in task che possono essere risolti attraverso rappresentazioni percettive, come giudizi di similarità visiva tra oggetti basati sulla loro apparenza percettiva, o matching di oggetti con le loro locazioni tipiche in scene visive.

Inversamente, pazienti con lesioni nelle cortecce sensoriali e motorie ma con preservazione relativa delle regioni linguistiche classiche dovrebbero mostrare il pattern opposto di deficit: difficoltà in task percettivi e immaginativi ma preservazione di capacità linguistiche simboliche. Evidenze empiriche da studi neuropsicologici forniscono supporto preliminare per questa predizione di doppia dissociazione. Pazienti con afasia semantica dovuta a lesioni nelle regioni temporali anteriori sinistre, che sono critiche per elaborazione semantica verbale, mostrano deficit sostanziali in task di denominazione e comprensione verbale ma possono mostrare performance sorprendentemente buona in task di matching visivo di oggetti e in task di uso funzionale di oggetti che possono essere risolti attraverso conoscenza percettiva e motoria delle affordances.

La seconda predizione riguarda effetti di concretezza nella memoria e nel processamento. Parole concrete con alto imageability, che attivano robustamente entrambi i sistemi verbale e immaginativo, dovrebbero mostrare vantaggi sostanziali rispetto a parole astratte con basso imageability, che attivano predominantemente soltanto il sistema verbale. Questi vantaggi dovrebbero manifestarsi in molteplici domini cognitivi. Nel dominio della memoria, l'effetto di concretezza predice che parole concrete dovrebbero essere ricordate significativamente meglio di parole astratte in task di memoria episodica come recall libero o riconoscimento, perché le parole concrete beneficiano di doppia codifica in entrambi i sistemi, creando tracce mnestiche ridondanti che aumentano la probabilità di recupero successivo.

Questa predizione è stata confermata attraverso letteralmente centinaia di studi sperimentali che hanno dimostrato effetti di concretezza robusti e replicabili nella memoria verbale. La magnitudine dell'effetto è tipicamente sostanziale, con differenze nell'accuratezza di recall che possono raggiungere venti a trenta punti percentuali a favore di parole concrete rispetto a parole astratte equate per frequenza, lunghezza e altre variabili lessicali. Inoltre, l'effetto di concretezza interagisce con il tipo di codifica: quando i partecipanti sono istruiti a usare strategie di immaginazione visiva durante la codifica, l'effetto di concretezza è amplificato, mentre quando sono istruiti a usare strategie di elaborazione verbale semantica, l'effetto è ridotto ma non eliminato completamente, consistente con l'idea che le parole concrete beneficiano di supporto immaginativo anche quando le strategie esplicite non lo incoraggiano.

Nel dominio del processamento online, l'effetto di concretezza predice che parole concrete dovrebbero essere processate più rapidamente di parole astratte in task di decisione lessicale, denominazione di immagini, e comprensione di frasi. Le evidenze empiriche confermano generalmente questa predizione, benché la direzione e magnitudine dell'effetto può variare con le richieste specifiche del task. In task di decisione lessicale dove i partecipanti devono giudicare rapidamente se una stringa di lettere è una parola vera o una non-parola, parole concrete tipicamente elicitano tempi di reazione più rapidi di parole astratte, con differenze medie nell'ordine di trenta a cinquanta millisecondi. Questo vantaggio di velocità è interpretato come riflettente l'attivazione parallela di rappresentazioni verbali e immaginative per parole concrete, con le attivazioni convergenti che accelerano la decisione lessicale.

La terza predizione riguarda modulazione contestualmente dipendente dell'attivazione delle componenti percettive e motorie. Se le rappresentazioni grounded sono componenti genuini e funzionalmente rilevanti del significato piuttosto che semplicemente associazioni epifenomenali post-comprensione, allora l'attivazione di queste rappresentazioni dovrebbe essere modulata sistematicamente dalle richieste del task e dal contesto pragmatico. Specificamente, in task che richiedono elaborazione semantica profonda del significato delle parole, l'attivazione delle cortecce sensoriali e motorie dovrebbe essere più forte rispetto a task che richiedono elaborazione superficiale basata soltanto su proprietà fonologiche o ortografiche.

Questa predizione è stata testata attraverso studi fMRI che hanno utilizzato paradigmi di livelli di elaborazione. In questi studi, i partecipanti processano le stesse parole in condizioni diverse che variano nella profondità di elaborazione semantica richiesta. In una condizione di elaborazione profonda, i partecipanti potrebbero essere istruiti a giudicare se ciascuna parola denota qualcosa che è piacevole o spiacevole, un giudizio che richiede accesso al significato semantico completo. In una condizione di elaborazione superficiale, i partecipanti potrebbero essere istruiti a giudicare se ciascuna parola contiene la lettera E, un giudizio che può essere fatto basandosi soltanto sulla forma ortografica senza accedere al significato.

I risultati rivelano che l'attivazione delle cortecce sensoriali e motorie durante la lettura di parole è significativamente maggiore nella condizione di elaborazione semantica profonda rispetto alla condizione di elaborazione superficiale, anche quando le parole stesse sono identiche attraverso le condizioni. Questo pattern supporta l'idea che l'attivazione grounded è modulata dalle richieste di elaborazione semantica e non è semplicemente un effetto automatico bottom-up della presentazione dello stimolo.

L'Integrazione dei Livelli Esplicativi: Dal Formale al Neurale

La Necessità di una Prospettiva Multi-Livello

Attraverso l'esplorazione sistematica condotta nei capitoli precedenti, abbiamo tracciato un percorso intellettuale che attraversa molteplici domini di indagine, dalla semantica formale alle neuroscienze cognitive, dalla pragmatica teorica agli studi sull'acquisizione. Ogni

dominio ha fornito intuizioni cruciali circa aspetti diversi del fenomeno complesso della comprensione linguistica, eppure ciascuno, preso isolatamente, cattura soltanto una frazione della realtà cognitiva che cerchiamo di comprendere. La semantica formale illumina la struttura compositiva e le relazioni inferenziali tra significati, ma rimane astratta rispetto all'implementazione psicologica effettiva. La pragmatica cognitiva cattura i processi inferenziali che arricchiscono il contenuto linguistico oltre il codificato letteralmente, ma spesso manca di precisione formale nelle sue formulazioni. Le neuroscienze rivelano i substrati neurali e le dinamiche temporali dell'elaborazione, ma la connessione tra oscillazioni neurali e operazioni semantiche rimane spesso speculativa.

La sfida teorica fondamentale che emerge è quella dell'integrazione: come possiamo costruire un framework unificato che rispetti i vincoli e incorpori le intuizioni di tutti questi livelli esplicativi, dall'analisi logica della struttura proposizionale fino ai pattern di sincronizzazione neuronale che la implementano nel tessuto biologico del cervello? Questa non è meramente una questione di completezza enciclopedica, ma tocca questioni profonde circa la natura stessa della spiegazione scientifica in scienze cognitive. Quando abbiamo veramente "spiegato" la comprensione linguistica? È sufficiente specificare le condizioni di verità degli enunciati, o dobbiamo anche specificare come quegli enunciati vengono processati in tempo reale? È sufficiente descrivere i processi computazionali astratti, o dobbiamo anche specificare come quei processi sono implementati in circuiti neurali? E se dobbiamo fare tutto questo, come assicuriamo coerenza tra i diversi livelli esplicativi?

La risposta che proponiamo è che una teoria adeguata della comprensione linguistica deve essere intrinsecamente multi-livello, operando simultaneamente a quello che David Marr, nel suo framework influente per la scienza della visione, chiamava i livelli computazionale, algoritmico e implementazionale. Al livello computazionale, specifichiamo cosa viene computato e perché: quali mappature da input linguistici a rappresentazioni semantiche il sistema deve implementare, e quali vincoli normativi o funzionali giustificano quella particolare mappatura. Al livello algoritmico, specifichiamo come la computazione viene eseguita: quali rappresentazioni interne e processi trasformazionali operano per produrre l'output desiderato dall'input. Al livello implementazionale, specifichiamo attraverso quale substrato fisico, in questo caso circuiti neurali biologici, l'algoritmo viene realizzato.

Il Livello Computazionale: Semantica Vero-Condizionale e Pragmatica Inferenziale

Al livello computazionale più astratto, la teoria della comprensione linguistica deve specificare la funzione che il sistema linguistico implementa. La tradizione della semantica vero-condizionale fornisce una specificazione precisa per almeno un aspetto di questa funzione: la comprensione di un enunciato dichiarativo consiste nell'identificazione delle condizioni sotto le quali quell'enunciato sarebbe vero. Formalmente, comprendere un enunciato S significa computare una funzione che mappa contesti, tipicamente modellati come tuple contenenti mondo possibile, tempo, parlante, ascoltatore, e altre dimensioni contestuali rilevanti, su valori di verità. Questa funzione cattura il contenuto proposizionale dell'enunciato, ciò che l'enunciato asserisce circa la realtà.

Tuttavia, come abbiamo visto estensivamente, questa caratterizzazione vero-condizionale, per quanto precisa, è drammaticamente sottodeterminata rispetto a ciò che viene

effettivamente comunicato attraverso gli enunciati. Il contenuto semantico codificato composizionalmente è soltanto uno scheletro schematico che richiede arricchimento pragmatico attraverso processi inferenziali per produrre l'interpretazione comunicata effettiva. La Teoria della Pertinenza fornisce una specificazione computazionale di questo processo di arricchimento: l'ascoltatore inferisce l'interpretazione che massimizza la pertinenza, definita come il rapporto tra effetti cognitivi positivi e sforzo di elaborazione, sotto il vincolo della presunzione di pertinenza ottimale comunicata ostensivamente dall'enunciato.

Possiamo quindi raffinare la specificazione computazionale: la funzione che il sistema linguistico implementa non mappa semplicemente enunciati su condizioni di verità, ma piuttosto mappa tuple di enunciati e contesti su interpretazioni che soddisfano criteri di ottimalità pragmatica. Formalmente: Comprensione colon insieme di Enunciato moltiplicato per Contesto verso Interpretazione tale che per ogni coppia di e e C, Comprensione(e, C) uguale a argmax rispetto a I nella formula Pertinenza(I, C) soggetto al vincolo che I è consistente con la presunzione di pertinenza ottimale comunicata da e.

Questa specificazione computazionale cattura il fatto normativo che la comprensione non è arbitraria ma è vincolata da principi di razionalità comunicativa: date certe assunzioni circa gli scopi cooperativi della comunicazione e i vincoli cognitivi degli interlocutori, certi interpretazioni sono ottimali mentre altre sono sub-ottimali. Il compito del sistema cognitivo è implementare un algoritmo che identifica in modo affidabile ed efficiente l'interpretazione ottimale.

Il Livello Algoritmico: Elaborazione Incrementale e Architetture Multi-Stadio

Scendendo dal livello computazionale astratto al livello algoritmico, dobbiamo specificare i processi attraverso cui l'interpretazione ottimale viene effettivamente derivata in tempo reale. Come abbiamo discusso, l'elaborazione linguistica è intrinsecamente incrementale: gli umani non aspettano di ricevere l'enunciato completo prima di iniziare l'interpretazione, ma costruiscono rappresentazioni interpretative gradualmente, parola per parola, man mano che l'input si dispiega temporalmente. Questa incrementalità impone vincoli architetturali significativi: il sistema deve mantenere rappresentazioni parziali che possono essere estese e revisionate man mano che nuova informazione arriva.

L'architettura algoritmica che abbiamo delineato è una pipeline multi-stadio con feedback bidirezionale. All'inizio della pipeline, il segnale acustico o ortografico grezzo viene trasformato in rappresentazioni fonologiche o grafematiche discrete attraverso processi di riconoscimento percettivo. Queste rappresentazioni fonologiche attivano poi voci lessicali nella memoria lessicale a lungo termine, recuperando informazione morfologica, sintattica, semantica, e pragmatica associata con quelle parole. Le rappresentazioni lessicali attivate vengono assemblate in strutture sintattiche attraverso processi di parsing che applicano conoscenze grammaticali per costruire alberi sintattici gerarchici. Le strutture sintattiche forniscono poi l'input per processi di interpretazione semantica compositiva che derivano rappresentazioni del significato attraverso applicazione di regole di composizione. Finalmente, queste rappresentazioni semantiche sottodeterminate vengono arricchite pragmaticamente attraverso processi inferenziali che integrano conoscenza contestuale e mondiale.

Crucialmente, questa architettura non è puramente feed-forward ma incorpora molteplici forme di feedback. L'informazione semantica e pragmatica influenza il parsing sintattico, aiutando a risolvere ambiguità strutturali attraverso vincoli di plausibilità. Le aspettative lessicali derivate dal contesto semantico modulano l'accesso lessicale, facilitando il riconoscimento di parole predicibili. Le rappresentazioni sintattiche parziali guidano l'attenzione percettiva verso informazione fonologica rilevante. Questo feedback multi-livello implementa un processo di elaborazione interattiva nel quale rappresentazioni a tutti i livelli si influenzano reciprocamente, convergendo verso un'interpretazione globalmente coerente.

Formalmente, possiamo modellare questa architettura interattiva come un sistema dinamico nel quale ciascun livello rappresentazionale è una variabile di stato che evolve temporalmente sotto l'influenza sia dell'input esterno che degli altri livelli. Sia R uguale a insieme di R_percezione, R_lessico, R_sintassi, R_semantica, R_pragmatica la tupla di stati rappresentazionali ai diversi livelli. L'evoluzione temporale è governata da un sistema di equazioni differenziali o differenze: dR_i diviso dt uguale a F_i(R, Input(t)), dove F_i specifica come lo stato al livello i cambia in funzione degli stati a tutti i livelli e dell'input corrente. Le funzioni F_i incorporano sia connessioni feed-forward che catturano il flusso di elaborazione bottom-up dall'input alla semantica, sia connessioni feedback che implementano vincoli top-down.

Questo sistema dinamico converge verso stati di equilibrio che corrispondono a interpretazioni coerenti attraverso tutti i livelli: stati nei quali le rappresentazioni percettive sono consistenti con le rappresentazioni lessicali attivate, che sono consistenti con le strutture sintattiche costruite, che sono consistenti con le interpretazioni semantiche e pragmatiche derivate. Il processo di convergenza può essere concepito come minimizzazione di una funzione di energia o errore globale che misura l'incoerenza tra livelli: Energia(R) uguale alla somma su tutti i livelli i,j di w_ij moltiplicato per Incoerenza(R_i, R_j), dove Incoerenza quantifica quanto gli stati ai livelli i e j sono mutuamente inconsistenti, e w_ij sono pesi che riflettono la forza delle connessioni tra livelli.

Il Livello Implementazionale: Dinamiche Oscillatorie e Binding Temporale

Al livello implementazionale più concreto, dobbiamo specificare come questi processi algoritmici sono realizzati nei circuiti neurali del cervello umano. Come abbiamo visto, le oscillazioni neurali in diverse bande di frequenza giocano ruoli cruciali nell'organizzazione temporale dell'elaborazione linguistica e nel coordinamento della comunicazione tra regioni distribuite. Le oscillazioni delta agganciano la struttura prosodica del discorso, fornendo un framework temporale per la segmentazione in unità frasali. Le oscillazioni theta organizzano l'elaborazione a livello di parole e sintagmi, supportando la memoria di lavoro sintattica. Le oscillazioni gamma implementano il binding locale di features in rappresentazioni coerenti e facilitano la comunicazione rapida tra regioni.

Crucialmente, queste oscillazioni non operano indipendentemente ma sono accoppiate gerarchicamente attraverso meccanismi di modulazione cross-frequenza. La fase delle oscillazioni lente modula l'ampiezza delle oscillazioni rapide, creando una cascata di modulazioni temporali che abbraccia tre ordini di grandezza: gamma nell'ordine di decine di millisecondi per features e morfemi, theta nell'ordine di centinaia di millisecondi per parole e sintagmi, delta nell'ordine di secondi per frasi ed enunciati. Questa architettura oscillatoria

gerarchica fornisce un meccanismo naturale per implementare la struttura gerarchica del linguaggio: elementi a diversi livelli della gerarchia linguistica occupano finestre temporali definite da oscillazioni a frequenze corrispondenti.

Il binding composizionale, il problema computazionale di assemblare rappresentazioni complesse mantenendo distinte molteplici strutture simultanee, può essere implementato attraverso sincronizzazione temporale nella banda gamma. Popolazioni neuronali che rappresentano elementi che devono essere legati insieme, come un nome e il suo ruolo tematico in una struttura argomentale, scaricano sincronicamente in una fase gamma comune, mentre popolazioni che rappresentano strutture distinte scaricano in fasi asincrone. La temporizzazione precisa degli spike nella scala dei millisecondi fornisce un codice temporale che specifica quali rappresentazioni appartengono insieme, risolvendo il problema del binding senza richiedere connessioni anatomiche dedicate tra tutte le possibili combinazioni di elementi.

Possiamo formalizzare questa implementazione neurale attraverso modelli di reti neurali oscillatorie. Rappresentiamo popolazioni neuronali attraverso oscillatori di fase accoppiati, dove ciascun oscillatore rappresenta un cluster di neuroni che codifica una feature o un elemento lessicale specifico. Lo stato di ciascun oscillatore è specificato dalla sua fase phi e dalla sua ampiezza A. L'evoluzione temporale è governata da equazioni che includono sia la dinamica intrinseca dell'oscillatore che i termini di accoppiamento con altri oscillatori: d_phi_i diviso dt uguale a omega_i più la somma su j di K_ij moltiplicato per il seno di phi_j meno phi_i, dove omega_i è la frequenza naturale dell'oscillatore i, e K_ij sono le forze di accoppiamento che possono essere positive per accoppiamento eccitatorio o negative per accoppiamento inibitorio.

Quando due oscillatori che devono essere legati insieme ricevono input coordinati, i termini di accoppiamento li portano a sincronizzarsi: le loro fasi convergono verso un valore comune. Oscillatori che rappresentano elementi che devono rimanere distinti, come argomenti di predicati diversi, possono essere accoppiati attraverso inibizione reciproca che mantiene le loro fasi separate. La rete nel suo insieme implementa quindi una forma di binding dinamico temporale dove le relazioni strutturali sono codificate attraverso pattern temporali di sincronizzazione e desincronizzazione.

Predizioni Testabili e Validazione Empirica del Framework Integrato

Un framework teorico, per quanto elegante e comprensivo, ha valore scientifico soltanto nella misura in cui genera predizioni empiriche testabili che possono essere confermate o falsificate attraverso osservazione e sperimentazione. Il framework multi-livello che abbiamo delineato genera una ricchezza di predizioni testabili che attraversano metodologie diverse, dalla psicolinguistica comportamentale alla neurofisiologia.

A livello comportamentale, il framework esibisce pattern specifici di facilitazione e interferenza in paradigmi di priming e compatibilità. Se la comprensione linguistica attiva genuinamente rappresentazioni percettivo-motorie grounded, allora dovremmo osservare che comprendere linguaggio che descrive azioni o percezioni in modalità specifiche facilita l'elaborazione in quelle modalità e interferisce con elaborazione incompatibile. Queste

predizioni sono state ampiamente confermate attraverso gli studi di compatibilità azione-frase e priming cross-modale che abbiamo discusso.

A livello neurale, il framework predice pattern specifici di attivazione cerebrale durante la comprensione linguistica. Concetti concreti che hanno forte contenuto percettivo-motorio dovrebbero attivare non solo reti semantiche transmodali ma anche cortecce sensoriali e motorie modality-specific. Verbi di azione dovrebbero attivare regioni motorie somatotopicamente organizzate corrispondenti agli effettori coinvolti. Queste predizioni sono state sistematicamente confermate attraverso studi di neuroimaging funzionale.

A livello elettrofisiologico, il framework prevede pattern specifici di oscillazioni neurali e accoppiamenti cross-frequenza durante l'elaborazione linguistica. Frasi sintatticamente complesse dovrebbero evocare incrementi di potenza theta e accoppiamento theta-gamma aumentato. La fase delle oscillazioni delta dovrebbe agganciarsi alla struttura prosodica del discorso. Violazioni semantiche dovrebbero evocare burst gamma transitori riflettenti risoluzione di conflitto. Molte di queste predizioni hanno ricevuto supporto preliminare, anche se questo rimane un'area attiva di ricerca con molte questioni aperte.

Crucialmente, il framework genera anche predizioni circa interazioni tra livelli che non sarebbero derivabili da teorie che operano a un singolo livello. Per esempio, predice che manipolazioni della sincronizzazione neurale attraverso stimolazione transcranica dovrebbero produrre effetti selettivi sulla comprensione linguistica dipendenti dalla frequenza: stimolazione delta dovrebbe influenzare la segmentazione prosodica, stimolazione theta dovrebbe influenzare la memoria di lavoro sintattica, stimolazione gamma dovrebbe influenzare il binding semantico. Predice che lesioni a regioni motorie dovrebbero causare deficit selettivi nella comprensione di verbi di azione corrispondenti. Predice che modulare il carico cognitivo dovrebbe influenzare quali aspetti del significato vengono elaborati: sotto carico elevato, l'elaborazione dovrebbe concentrarsi su contenuto simbolico essenziale, mentre sotto carico ridotto può incorporare simulazioni percettivo-motorie ricche.

Questioni Aperte e Direzioni di Ricerca Futura

Nonostante i progressi sostanziali che abbiamo documentato, molteplici questioni fondamentali rimangono aperte e costituiscono direzioni promettenti per ricerca futura. Una questione centrale riguarda la natura precisa della relazione tra rappresentazioni simboliche e grounded. Abbiamo proposto un'architettura dual-coding nella quale i due tipi di rappresentazioni coesistono e interagiscono, ma i dettagli di come questa interazione opera rimangono largamente non specificati. Quali sono i meccanismi neurali attraverso cui l'attivazione simbolica si propaga alle cortecce sensoriali e motorie? Come viene controllato il grado di attivazione grounded, dato che evidentemente non ogni parola compresa evoca simulazioni percettive elaborate? Quali fattori contestuali e strategici modulano l'utilizzo relativo di elaborazione simbolica versus grounded?

Una seconda questione riguarda la scalabilità dei meccanismi di binding temporale attraverso sincronizzazione gamma. Mentre questa proposta è elegante e ha ricevuto supporto empirico per casi relativamente semplici, rimane non chiaro se può scalare a rappresentazioni linguistiche della complessità che caratterizza il linguaggio naturale. Frasi con annidamento sintattico multiplo e molteplici clausole subordinate richiedono il

mantenimento simultaneo di numerose strutture argomentali distinte. Se ciascuna struttura richiede una fase gamma distinta, e se la risoluzione di fase è limitata, potrebbero esserci vincoli fondamentali sulla complessità rappresentazionale supportabile attraverso binding temporale. Ricerca futura potrebbe esplorare se esistono meccanismi addizionali, oltre la separazione di fase, che contribuiscono al binding in strutture complesse.

Una terza questione concerne la generalizzabilità cross-linguistica del framework. La maggior parte della ricerca empirica che abbiamo discusso è stata condotta con parlanti di lingue indoeuropee, particolarmente inglese. Tuttavia, le lingue del mondo variano dramaticamente nelle loro proprietà strutturali: alcune sono prevalentemente configurazionali con ordine di parole rigido, altre sono non-configurazionali con ordine libero; alcune sono prevalentemente analitiche con morfologia scarsa, altre sono sintetiche o polisintetiche con morfologia ricca; alcune hanno struttura tonale dove pitch lessicale è fonologicamente distintivo, altre non hanno tono lessicale. Queste differenze strutturali potrebbero avere conseguenze per i meccanismi neurali di elaborazione. Le oscillazioni delta si agganciano alla prosodia in lingue tonali come fanno in lingue non-tonali? La memoria di lavoro theta opera diversamente in lingue con morfologia ricca dove parole singole possono contenere informazione equivalente a frasi intere in lingue analitiche? Studi cross-linguistici sistematici sono necessari per determinare quanto i meccanismi neurali dell'elaborazione linguistica sono universali versus dipendenti dalla lingua specifica.

Implicazioni Filosofiche: Significato, Mente e Realtà

Concludiamo questa esplorazione considerando alcune implicazioni filosofiche più ampie che emergono dal framework integrato che abbiamo sviluppato. La comprensione linguistica, vista attraverso questa lente multi-livello, non è un processo di semplice decodifica di significati pre-esistenti, ma è piuttosto un processo costruttivo complesso che integra informazione codificata linguisticamente con conoscenza contestuale, aspettative pragmatiche, e simulazioni percettivo-motorie, producendo rappresentazioni interpretative ricche che vanno drammaticamente oltre ciò che è letteralmente detto.

Questa concezione costruttiva della comprensione ha ramificazioni per questioni filosofiche classiche circa la natura del significato linguistico. Se il significato non è semplicemente codificato nelle espressioni linguistiche ma emerge dall'interazione tra forma linguistica, contesto, e conoscenze del comprendente, allora la distinzione netta tra semantica e pragmatica diventa problematica. Non c'è un livello di significato "letterale" completamente determinato dalla forma linguistica indipendentemente dal contesto; piuttosto, ci sono gradi di dipendenza contestuale, con alcuni aspetti del contenuto più stabili attraverso contesti e altri più variabili.

Inoltre, l'ancoraggio del significato linguistico in rappresentazioni percettivo-motorie grounded suggerisce una forma di embodied realism che media tra realismo metafisico tradizionale e relativismo concettuale. I nostri concetti non sono arbitrari o puramente convenzionali, ma sono vincolati dalla struttura della nostra esperienza corporea e percettiva nel mondo. Allo stesso tempo, non sono mere copie passive di una realtà indipendente dalla mente, ma sono costruzioni attive che riflettono le modalità specifiche attraverso cui i nostri sistemi sensoriali e motori interagiscono con l'ambiente. La realtà che rappresentiamo

attraverso il linguaggio è sempre già filtrata attraverso le lenti della nostra incarnazione biologica.

Infine, la prospettiva multi-livello che abbiamo adottato ha implicazioni per questioni circa le relazioni tra mente e cervello. L'elaborazione linguistica emerge dall'orchestrazione coordinata di molteplici sistemi neurali operanti a scale temporali e spaziali diverse. Non c'è un singolo "centro linguistico" ma piuttosto reti distribuite dinamicamente configurate. Le proprietà computazionali dell'elaborazione linguistica, come la composizionalità e il binding strutturale, non sono misteriose proprietà emergenti completamente disgiunte dal substrato neurale, ma possono essere comprese come risultanti da principi organizzativi specifici dell'architettura neurale, particolarmente la dinamica oscillatoria e la plasticità sinaptica. La mente linguistica è inseparabile dal cervello che la implementa, eppure le descrizioni a livelli diversi, computazionale versus implementazionale, catturano aspetti complementari della realtà cognitiva che non sono mutuamente riducibili.

Formalizzazione Matematica Rigorosa dei Meccanismi Neurali della Comprensione Linguistica

Parte I: Dinamiche Oscillatorie e Teoria dei Sistemi Accoppiati

Fondamenti Matematici delle Oscillazioni Neurali

Per sviluppare una teoria matematicamente rigorosa delle oscillazioni neurali come substrato computazionale della comprensione linguistica, dobbiamo innanzitutto stabilire il formalismo attraverso cui rappresentare l'attività oscillatoria di popolazioni neuronali. Consideriamo una popolazione neuronale caratterizzata da un campo di potenziale locale mediato nel tempo, che possiamo rappresentare attraverso una funzione V(t) del tempo. Questo potenziale riflette la somma pesata dell'attività elettrica post-sinaptica di migliaia di neuroni organizzati in architetture colonnari nella corteccia cerebrale.

La componente oscillatoria di questo segnale può essere estratta attraverso decomposizione spettrale. Applicando la trasformata di Fourier al segnale temporale, otteniamo una rappresentazione nel dominio della frequenza che ci permette di identificare le componenti ritmiche dominanti. Formalmente, la trasformata di Fourier del segnale V(t) è definita come $\hat{V}(f) = \int_{-\infty}^{+\infty} V(t) \exp(-2\pi i f t)\, dt$, dove f rappresenta la frequenza in hertz e i è l'unità immaginaria. Il modulo al quadrato di questa trasformata, $|\hat{V}(f)|^2$, fornisce lo spettro di potenza che quantifica l'energia del segnale presente in ciascuna banda di frequenza.

Tuttavia, per catturare la dinamica temporale delle oscillazioni, che è cruciale per comprendere come le oscillazioni si sincronizzano con strutture linguistiche temporalmente estese, necessitiamo di una rappresentazione tempo-frequenza. La trasformata wavelet continua fornisce precisamente questa rappresentazione. Definiamo la trasformata wavelet di V(t) rispetto a una wavelet madre ψ(t) come $W(a,b) = (1/\sqrt{a}) \int_{-\infty}^{+\infty} V(t)\, \psi^*((t-b)/a)\, dt$, dove a è il parametro di scala che controlla la frequenza della wavelet, b è il parametro di

traslazione temporale che determina la localizzazione temporale, e ψ* indica il complesso coniugato della wavelet madre.

Una scelta particolarmente appropriata per l'analisi di segnali oscillatori neurali è la wavelet di Morlet, definita come $\psi(t) = \pi^{-1/4} \exp(i\omega_0 t) \exp(-t^2/2)$, dove ω_0 è la frequenza centrale della wavelet, tipicamente scelta come $\omega_0 = 6$ per ottenere un buon bilanciamento tra risoluzione temporale e risoluzione frequenziale. La trasformata wavelet con wavelet di Morlet produce uno scalogramma che rappresenta l'energia del segnale come funzione sia del tempo che della frequenza, permettendoci di tracciare come la potenza nelle diverse bande oscillatorie varia dinamicamente durante l'elaborazione linguistica.

Modellazione della Fase Oscillatoria e Sincronizzazione

Per comprendere i meccanismi di sincronizzazione oscillatoria che implementano il binding temporale, dobbiamo rappresentare esplicitamente la fase delle oscillazioni. Dato un segnale oscillatorio V(t), possiamo estrarre la sua componente di fase istantanea φ(t) attraverso la trasformata di Hilbert. La trasformata di Hilbert di V(t) è definita come $HV = (1/\pi)\ P.V.\ \int_{-\infty}^{+\infty} V(\tau)/(t-\tau)\ d\tau$, dove P.V. denota il valore principale di Cauchy dell'integrale. Costruiamo poi il segnale analitico come $Z(t) = V(t) + iHV = A(t)\exp(i\varphi(t))$, dove $A(t) = \sqrt{(V^2(t) + H[V]^2(t))}$ è l'ampiezza istantanea e $\varphi(t) = \arctan(HV/V(t))$ è la fase istantanea.

La fase istantanea φ(t) è una variabile angolare che evolve continuamente, tipicamente in modo approssimativamente lineare per oscillazioni regolari: $\varphi(t) \approx \omega t + \varphi_0$, dove ω è la frequenza angolare istantanea e φ_0 è un offset di fase. La frequenza istantanea può essere estratta come $f(t) = (1/2\pi)(d\varphi/dt)$. Questa rappresentazione di fase è fondamentale perché la sincronizzazione tra oscillazioni in diverse popolazioni neuronali è definita attraverso relazioni di fase stabile.

Consideriamo ora due popolazioni neuronali, designate come popolazione uno e popolazione due, ciascuna caratterizzata da un'oscillazione con fase $\varphi_1(t)$ e $\varphi_2(t)$ rispettivamente. La sincronizzazione di fase tra queste popolazioni è quantificata attraverso la differenza di fase $\Delta\varphi(t) = \varphi_1(t) - \varphi_2(t)$. Se le popolazioni sono perfettamente sincronizzate, questa differenza di fase rimane costante nel tempo: Δφ(t) = costante. Se invece le oscillazioni sono completamente asincrone, la differenza di fase vaga liberamente attraverso tutto l'intervallo da zero a due pi greco.

Per misurare quantitativamente il grado di sincronizzazione in un ensemble di prove sperimentali, calcoliamo il Phase Locking Value, definito come $PLV = |\langle \exp(i\Delta\varphi_n) \rangle_n|$, dove la media angolare $\langle \cdot \rangle_n$ è calcolata attraverso N prove sperimentali e $\Delta\varphi_n$ è la differenza di fase nella prova n-esima. Il PLV varia da zero, indicando assenza completa di sincronizzazione con fasi distribuite uniformemente, a uno, indicando sincronizzazione perfetta con differenza di fase costante attraverso tutte le prove.

Equazioni Differenziali per Oscillatori Accoppiati

Per comprendere come la sincronizzazione emerge dinamicamente dall'interazione tra popolazioni neuronali, dobbiamo formulare equazioni differenziali che governano l'evoluzione temporale delle fasi oscillatorie. Il modello più semplice e ampiamente utilizzato

è il modello di Kuramoto, che descrive un sistema di oscillatori accoppiati attraverso le loro fasi.

Consideriamo un sistema di N oscillatori, ciascuno con fase $\varphi_i(t)$ e frequenza naturale ω_i. L'evoluzione temporale della fase dell'oscillatore i-esimo è governata dall'equazione differenziale $d\varphi_i/dt = \omega_i + (K/N) \Sigma_{j=1}^{N} \sin(\varphi_j - \varphi_i)$, dove K è la forza di accoppiamento che determina quanto fortemente gli oscillatori si influenzano reciprocamente. Il termine $\sin(\varphi_j - \varphi_i)$ implementa un accoppiamento che dipende dalla differenza di fase: quando l'oscillatore j è in anticipo rispetto all'oscillatore i ($\varphi_j > \varphi_i$), il termine di accoppiamento è positivo e accelera l'oscillatore i; quando è in ritardo, il termine è negativo e rallenta l'oscillatore i. Questo meccanismo di feedback produce convergenza verso sincronizzazione.

La dinamica collettiva del sistema può essere caratterizzata attraverso il parametro d'ordine complesso $r(t)\exp(i\Psi(t)) = (1/N) \Sigma_{j=1}^{N} \exp(i\varphi_j(t))$, dove r(t) misura il grado di coerenza di fase dell'ensemble: r = 0 indica completa incoerenza con fasi distribuite uniformemente, mentre r = 1 indica sincronizzazione perfetta con tutte le fasi identiche. L'angolo $\Psi(t)$ rappresenta la fase media del sistema.

Per il caso speciale di oscillatori identici con $\omega_i = \omega$ per tutti i, il sistema di Kuramoto può essere riscritto in forma più compatta come $d\varphi_i/dt = \omega + Kr \sin(\Psi - \varphi_i)$. Questa formulazione rivela che ciascun oscillatore è effettivamente accoppiato a un campo medio generato dall'ensemble, con forza proporzionale a Kr. Esiste una transizione di fase critica nel comportamento del sistema: per valori della forza di accoppiamento inferiori a una soglia critica $K < Kc$, il sistema rimane incoerente con $r \approx 0$; per $K > Kc$, emerge spontaneamente sincronizzazione parziale con $r > 0$ che aumenta progressivamente con l'aumentare di K.

Per oscillatori con frequenze naturali distribuite secondo una densità di probabilità $g(\omega)$, la transizione di sincronizzazione può essere caratterizzata analiticamente. Nel limite termodinamico $N \to \infty$, si può dimostrare che il parametro d'ordine soddisfa l'equazione di autoconsistenza $r = \int d\omega\, g(\omega) \int_0^{2\pi} (d\theta/2\pi) |d\varphi/dt|^{-1}$, dove $\varphi(\omega,t)$ è la soluzione dell'equazione di Kuramoto per un oscillatore con frequenza naturale ω.

Accoppiamento Cross-Frequenza e Modulazione di Ampiezza

I meccanismi di sincronizzazione che abbiamo descritto operano tra oscillazioni nella stessa banda di frequenza. Tuttavia, come abbiamo discusso nel contesto delle gerarchie temporali nell'elaborazione linguistica, un aspetto cruciale dell'organizzazione neurale è l'accoppiamento tra oscillazioni a frequenze diverse, particolarmente la modulazione dell'ampiezza di oscillazioni rapide attraverso la fase di oscillazioni lente.

Formalizziamo questo fenomeno considerando due segnali oscillatori: un'oscillazione lenta a bassa frequenza $V_{low}(t)$ con fase $\varphi_{low}(t)$ e un'oscillazione rapida ad alta frequenza $V_{high}(t)$ con ampiezza istantanea $A_{high}(t)$. L'accoppiamento fase-ampiezza, o Phase-Amplitude Coupling, si verifica quando l'ampiezza dell'oscillazione rapida varia sistematicamente in funzione della fase dell'oscillazione lenta.

Possiamo modellare questo accoppiamento attraverso una funzione di modulazione che specifica come l'ampiezza dipende dalla fase: $A_{high}(t) = A_0(1 + m \cdot h(\varphi_{low}(t)))$, dove A_0 è l'ampiezza baseline, m è l'indice di modulazione che quantifica la forza dell'accoppiamento

(con m = 0 indicando assenza di modulazione e m → 1 indicando modulazione forte), e h(φ) è una funzione periodica che specifica il pattern di modulazione. Una scelta naturale per h(φ) è la funzione cosinusoidale centrata sulla fase preferita: h(φ) = cos(φ - φ□ref), dove φ□ref è la fase dell'oscillazione lenta alla quale l'ampiezza dell'oscillazione rapida è massimale.

Per quantificare empiricamente il grado di accoppiamento fase-ampiezza presente nei dati neurali, calcoliamo il Modulation Index attraverso il seguente procedimento. Primo, estraiamo la fase dell'oscillazione lenta φ□ow(t) e l'ampiezza dell'oscillazione rapida A□ig□(t) attraverso filtraggio passa-banda nelle rispettive bande di frequenza seguito da trasformata di Hilbert. Secondo, discretizziamo lo spazio delle fasi in b bins, tipicamente b = 18 bins corrispondenti a intervalli di venti gradi. Terzo, per ciascun bin j, calcoliamo l'ampiezza media dell'oscillazione rapida quando la fase lenta cade in quel bin: ⟨A□ig□⟩□. Quarto, normalizziamo questi valori per ottenere una distribuzione di probabilità: p□ = ⟨A□ig□⟩□ / Σ□₌₁ᵇ ⟨A□ig□⟩□.

Se non c'è accoppiamento, questa distribuzione sarà uniforme: p□ = 1/b per tutti i j. Se c'è forte accoppiamento, la distribuzione sarà concentrata attorno a certe fasi preferite. Quantifichiamo la deviazione dall'uniformità attraverso l'entropia di Kullback-Leibler: MI = (log(b) - H(p)) / log(b), dove H(p) = -Σ□₌₁ᵇ p□ log(p□) è l'entropia di Shannon della distribuzione. Il Modulation Index varia da MI = 0 per distribuzione perfettamente uniforme (nessun accoppiamento) a MI = 1 per distribuzione perfettamente concentrata su un singolo bin (accoppiamento massimo).

Modelli Neurali di Generazione di Oscillazioni

Per comprendere come le oscillazioni emergono dalle proprietà biofisiche delle reti neuronali, dobbiamo formulare modelli che incorporano esplicitamente le dinamiche dei neuroni individuali e della loro interconnettività sinaptica. Un approccio standard è rappresentare ciascun neurone attraverso un modello a compartimento singolo che cattura le proprietà elettriche essenziali della membrana cellulare.

Il modello neuronale integratore a perdita, o leaky integrate-and-fire, descrive l'evoluzione temporale del potenziale di membrana V(t) di un neurone attraverso l'equazione differenziale τ□ dV/dt = -(V - V$_{re}$□□) + RI□$_{\gamma}$□(t), dove τ□ è la costante di tempo della membrana, tipicamente nell'ordine di dieci-venti millisecondi per neuroni corticali, V$_{re}$□□ è il potenziale di riposo, R è la resistenza di input della membrana, e I□$_{\gamma}$□(t) è la corrente sinaptica totale ricevuta dal neurone. Quando il potenziale di membrana raggiunge una soglia V□□$_{re}$□□, il neurone emette uno spike e il potenziale viene resettato a un valore V$_{re}$□$_{e}$□ < V$_{re}$□□, dopo il quale rimane refrattario per un periodo τ$_{re}$f.

La corrente sinaptica I□$_{\gamma}$□(t) è la somma dei contributi da tutti i neuroni presinaptici che proiettano sul neurone in questione: I□$_{\gamma}$□(t) = Σ□ w□ Σ□ α(t - t□ᵏ), dove w□ è il peso sinaptico della connessione dal neurone presinaptico j, t□ᵏ sono i tempi degli spike emessi dal neurone j, e α(t) è la funzione kernel post-sinaptico che descrive il decorso temporale della corrente evocata da un singolo spike presinaptico. Una forma comune per questa funzione kernel è la alfa-function: α(t) = (t/τ□$_{\gamma}$□) exp(-t/τ□$_{\gamma}$□) per t ≥ 0, dove τ□$_{\gamma}$□ è la costante di tempo sinaptica, tipicamente nell'ordine di pochi millisecondi per sinapsi veloci.

Per generare oscillazioni collettive, la rete deve incorporare meccanismi di feedback. Un circuito canonico per la generazione di oscillazioni gamma consiste in una popolazione di neuroni eccitatori reciprocamente connessi e una popolazione di neuroni inibitori, anche reciprocamente connessi, con connessioni eccitatorie dai neuroni eccitatori agli inibitori e connessioni inibitorie dagli inibitori agli eccitatori. Questo circuito, chiamato loop eccitatorio-inibitorio, genera oscillazioni attraverso il seguente meccanismo: un aumento dell'attività eccitatoria recluta neuroni inibitori, che producono inibizione ricorrente che sopprime l'attività eccitatoria; la decadenza dell'inibizione permette poi la riattivazione dell'eccitazione, chiudendo il ciclo.

La frequenza delle oscillazioni generate da questo circuito dipende dai ritardi sinaptici e dalle costanti di tempo delle correnti sinaptiche inibitorie. Formalmente, considerando sinapsi inibitorie mediate da recettori GABA-A con costante di tempo τ_{ib} e un ritardo di conduzione sinaptico δ, la frequenza di oscillazione è approssimativamente $f \approx 1 / (2(\tau_{ib} + \delta))$. Per $\tau_{ib} \approx 10$ millisecondi e $\delta \approx 2$ millisecondi, otteniamo $f \approx 42$ hertz, nel mezzo della banda gamma.

Modelli di Campo Medio per Dinamiche di Popolazione

Simulare esplicitamente le dinamiche di migliaia di neuroni interconnessi è computazionalmente oneroso. Un approccio alternativo è formulare equazioni di campo medio che descrivono le dinamiche aggregate della popolazione, ignorando le fluttuazioni a livello di singolo neurone. Questi modelli operano al livello di variabili macroscopiche come il tasso di firing medio della popolazione.

Consideriamo una popolazione omogenea di N neuroni integrators a perdita, ciascuno ricevente input sinaptici sia da neuroni all'interno della popolazione che da sorgenti esterne. Nel limite di grande popolazione $N \to \infty$, la dinamica della popolazione può essere descritta attraverso la densità di probabilità $\rho(V,t)$ che specifica la frazione di neuroni che hanno potenziale di membrana V al tempo t. L'evoluzione di questa densità è governata dall'equazione di Fokker-Planck: $\partial\rho/\partial t = -\partial/\partial V [(f(V) + \mu(t))\rho] + (\sigma^2/2) \partial^2\rho/\partial V^2$, dove $f(V) = -(V - V_{re})/\tau$ è la parte deterministica del drift, $\mu(t)$ è l'input sinaptico medio alla popolazione, e σ^2 caratterizza la varianza dell'input dovuta a fluttuazioni stocastiche.

Il tasso di firing della popolazione è dato dal flusso di probabilità attraverso la soglia: $r(t) = -\int dV\ J(V,t)|_{\{V=V_{re}\}}$, *dove* $J(V,t) = [f(V) + \mu(t)]\rho(V,t) - (\sigma^2/2) \partial\rho/\partial V$ *è la corrente di probabilità. Per input costanti in regime stazionario, la distribuzione di potenziale converge a uno stato stazionario e il tasso di firing diventa costante, determinato implicitamente attraverso la condizione di normalizzazione* $\int_{V_{re}}^{V_{re}} \rho(V)\, dV + \tau_{ref} r = 1$.

Un modello di campo medio ancora più ridotto rappresenta la popolazione semplicemente attraverso il suo tasso di firing medio r(t), con dinamica governata da un'equazione differenziale del tipo $\tau_o\, dr/dt = -r + F(I(t))$, dove τ_o è una costante di tempo effettiva della popolazione e F è una funzione di trasferimento che mappa l'input corrente totale I(t) sul tasso di firing stazionario. Una scelta comune per F è la funzione sigmoide: $F(I) = r_{ax} / (1 + \exp(-(I - I_{re})/\Delta I))$, dove r_{ax} è il tasso di firing massimo, I_{re} è la soglia di input, e ΔI determina la pendenza della curva.

Per una rete di popolazioni interconnesse, otteniamo un sistema di equazioni differenziali ordinarie accoppiate: $\tau_i \, dr_i/dt = -r_i + F_i(\Sigma_j \, w_{ij} \, r_j + I_{ext,i})$, dove l'indice i scorre sulle popolazioni, w_{ij} rappresenta la forza della connessione sinaptica dalla popolazione j alla popolazione i, e $I_{ext,i}$ è l'input esterno alla popolazione i. Questo sistema può esibire comportamento oscillatorio quando i pesi sinaptici e le costanti di tempo sono appropriatamente configurati per generare feedback negativo ritardato.

Parte II: Modelli Bayesiani e Teoria dell'Informazione nella Comprensione

Inferenza Bayesiana come Framework Computazionale

La comprensione linguistica, come abbiamo argomentato, è fondamentalmente un processo inferenziale nel quale l'ascoltatore deve recuperare le intenzioni comunicative del parlante basandosi su evidenza linguistica sottodeterminata. Questo processo può essere formalizzato con precisione attraverso il framework dell'inferenza bayesiana, che fornisce principi normativi per il ragionamento sotto incertezza.

Sia H lo spazio delle possibili interpretazioni o ipotesi che l'ascoltatore considera, e sia D l'evidenza osservata, consistente nell'enunciato linguistico e nel contesto percettivo e discorsivo. Il teorema di Bayes specifica come l'ascoltatore dovrebbe aggiornare le sue credenze circa le interpretazioni alla luce dell'evidenza: $P(H|D) = P(D|H)P(H) / P(D)$, dove $P(H|D)$ è la distribuzione di probabilità posteriore sulle interpretazioni dato l'evidenza, $P(D|H)$ è la verosimiglianza dell'evidenza sotto l'ipotesi H, $P(H)$ è la distribuzione di probabilità a priori sulle interpretazioni prima di osservare l'evidenza, e $P(D) = \Sigma_h \, P(D|h)P(h)$ è la probabilità marginale dell'evidenza, ottenuta sommando su tutte le possibili interpretazioni.

La verosimiglianza $P(D|H)$ cattura il modello generativo attraverso cui le interpretazioni producono evidenza linguistica osservabile. Nel contesto della comprensione linguistica, questo corrisponde a un modello di produzione: dato che il parlante intende comunicare l'interpretazione H, con quale probabilità produrrebbe l'enunciato osservato D? Questo modello deve incorporare vincoli grammaticali che rendono certe formulazioni linguistiche più o meno probabili per esprimere certi contenuti, nonché considerazioni pragmatiche circa l'efficienza comunicativa che favoriscono espressioni più concise e meno ambigue quando possibile.

Il prior $P(H)$ incorpora conoscenze pre-esistenti e bias dell'ascoltatore circa quali interpretazioni sono più o meno plausibili indipendentemente dall'evidenza linguistica specifica. Questo può includere conoscenze circa la frequenza relativa di diverse strutture semantiche nel linguaggio, aspettative derivate dal contesto discorsivo precedente, e vincoli di plausibilità derivati dalla conoscenza del mondo. Per esempio, un'interpretazione che attribuisce intenzioni comunicative bizzarre o irrazionali al parlante riceverà probabilità a priori bassa, riflettendo l'assunzione cooperativa che i parlanti agiscono in modo razionale per raggiungere scopi comunicativi.

L'interpretazione ottimale secondo il criterio bayesiano è quella che massimizza la probabilità posteriore, chiamata stima a massimo a posteriori o MAP: $H_{MAP} = \text{argmax}_H$

P(H|D) = argmax_H P(D|H)P(H), dove l'ultima uguaglianza segue dal fatto che P(D) è costante rispetto a H e può essere ignorato nell'ottimizzazione.

Formalizzazione del Rational Speech Act Model

Un'applicazione particolarmente influente dell'inferenza bayesiana alla pragmatica è il Rational Speech Act model sviluppato da Noah Goodman e collaboratori, che fornisce una teoria computazionale precisa di come gli ascoltatori inferiscono significati pragmaticamente arricchiti attraverso ragionamento circa parlanti razionali.

Il modello definisce una gerarchia di agenti cognitivi che ragionano ricorsivamente l'uno circa l'altro. Al livello più basso, definiamo un parlante letterale L_0 che produce enunciati proporzionalmente alla loro verità letterale: $P\_L_0(u|s) \propto \exp(\alpha\ [[u]]^s)$, dove u è un enunciato, s è uno stato del mondo, [[u]]^s è la funzione caratteristica che vale uno se u è letteralmente vero in s e zero altrimenti, e α è un parametro di razionalità che controlla quanto deterministicamente il parlante sceglie enunciati veri. Nel limite $\alpha \to \infty$, il parlante letterale produce soltanto enunciati letteralmente veri con probabilità uniforme; per α finito, può occasionalmente produrre enunciati falsi con probabilità decrescente.

Un ascoltatore pragmatico P_1 inferisce lo stato del mondo assumendo che il parlante sia un parlante letterale L_0: $P\_P_1(s|u) \propto P\_L_0(u|s)\ P(s)$, dove P(s) è il prior sugli stati del mondo. Applicando il teorema di Bayes e sostituendo la definizione di P_L_0, otteniamo $P\_P_1(s|u) \propto \exp(\alpha\ [[u]]^s)\ P(s)$. Questo ascoltatore inferisce che stati dove l'enunciato è letteralmente vero sono più probabili, con il peso modulato dal prior.

Un parlante pragmatico S_1 produce enunciati scegliendo quello che massimizza l'utilità comunicativa, definita come la probabilità che un ascoltatore pragmatico P_1 inferirà correttamente lo stato inteso: $P\_S_1(u|s) \propto \exp(\beta\ (\log P\_P_1(s|u) - C(u)))$, dove β è un parametro di razionalità del parlante e C(u) è un costo che penalizza enunciati più lunghi o complessi, implementando una preferenza per concisione. Espandendo $\log P\_P_1(s|u)$ usando la regola di Bayes, otteniamo $P\_S_1(u|s) \propto \exp(\beta\ (\log P\_L_0(u|s) + \log P(s) - \log P(u) - C(u)))$.

Finalmente, un ascoltatore pragmatico P_2 inferisce lo stato assumendo che il parlante sia un parlante pragmatico S_1: $P\_P_2(s|u) \propto P\_S_1(u|s)\ P(s)$. Questa gerarchia può essere estesa a livelli arbitrari di ricorsività, benché evidenze empiriche suggeriscano che la maggior parte del ragionamento pragmatico umano opera ai livelli P_1 o P_2.

Questo framework genera predizioni quantitative precise circa fenomeni pragmatici come le implicature scalari. Consideriamo un contesto dove ci sono tre livelli possibili di una scala, per esempio nessuno, alcuni, tutti. Un parlante che osserva lo stato "tutti" potrebbe produrre l'enunciato "tutti" o l'enunciato più debole "alcuni" (che è letteralmente vero anche quando tutti sono il caso). Il modello RSA predice che il parlante pragmatico S_1 preferirà "tutti" perché è più informativo e permette all'ascoltatore di inferire lo stato corretto con maggiore certezza. Conseguentemente, quando l'ascoltatore pragmatico P_2 sente "alcuni", inferisce attraverso ragionamento alternativo che lo stato non può essere "tutti" (altrimenti il parlante avrebbe detto "tutti"), producendo l'implicatura scalare "alcuni ma non tutti".

Teoria dell'Informazione e Efficienza Comunicativa

Un aspetto complementare della formalizzazione bayesiana della comprensione è la prospettiva informazionale, che quantifica la comunicazione linguistica attraverso concetti della teoria dell'informazione di Shannon. La quantità fondamentale è l'entropia di Shannon, che misura l'incertezza media associata a una variabile casuale.

Data una variabile casuale discreta X con distribuzione di probabilità P(x), l'entropia è definita come $H(X) = -\Sigma_x P(x) \log_2 P(x)$, misurata in bits. L'entropia è massimale quando la distribuzione è uniforme, riflettendo massima incertezza, e minimale (uguale a zero) quando la distribuzione è concentrata su un singolo valore, riflettendo certezza completa. Per una variabile con n possibili valori equiprobabili, l'entropia è $H(X) = \log_2 n$ bits.

L'informazione mutua tra due variabili casuali X e Y quantifica quanto l'osservazione di una variabile riduce l'incertezza circa l'altra: $I(X;Y) = \Sigma_x \Sigma_y P(x,y) \log_2 (P(x,y) / (P(x)P(y)))$. Questa quantità è simmetrica, $I(X;Y) = I(Y;X)$, e può essere espressa come la riduzione di entropia: $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$, dove $H(X|Y) = -\Sigma_x \Sigma_y P(x,y) \log_2 P(x|y)$ è l'entropia condizionale che misura l'incertezza residua circa X dopo aver osservato Y.

Nel contesto della comunicazione linguistica, possiamo concepire il processo come un canale comunicativo che trasferisce informazione dallo stato del mondo o dall'intenzione comunicativa del parlante (la sorgente) all'interpretazione derivata dall'ascoltatore (il ricevente) attraverso il mezzo dell'enunciato linguistico. L'efficienza del canale è caratterizzata dalla rate di informazione mutua tra sorgente e interpretazione.

Formalmente, sia S la variabile casuale rappresentante lo stato inteso dal parlante, sia U l'enunciato prodotto, e sia $\hat{S}$ lo stato inferito dall'ascoltatore. La catena comunicativa è $S \rightarrow U \rightarrow \hat{S}$. La capacità informazionale del canale linguistico è $I(S;\hat{S}) = H(S) - H(S|\hat{S})$, che misura quanto dell'incertezza originale circa lo stato viene risolta dall'interpretazione dell'ascoltatore. Comunicazione perfetta corrisponderebbe a $I(S;\hat{S}) = H(S)$, ovvero completa risoluzione dell'incertezza.

Tuttavia, vincoli cognitivi e ambiguità inerenti al linguaggio implicano che tipicamente $I(S;\hat{S}) < H(S)$: c'è perdita informazionale nel canale. Questa perdita può essere decomposta attraverso la data processing inequality: $I(S;\hat{S}) \leq I(S;U)$, con uguaglianza se e solo se $\hat{S}$ è una funzione deterministica di U (ovvero, l'interpretazione è completamente determinata dall'enunciato senza ambiguità). La disuguaglianza riflette il fatto che il processo di interpretazione dell'ascoltatore non può recuperare informazione che non era presente nell'enunciato.

Codifica Efficiente e Trade-off Complessità-Precisione

La teoria dell'informazione fornisce anche un framework per comprendere come le lingue naturali bilanciano trade-offs tra espressività e complessità. Il principio di codifica efficiente suggerisce che le lingue dovrebbero allocare risorse codificanti (lunghezza di espressioni, complessità morfo-sintattica) in modo inversamente proporzionale alla frequenza o probabilità dei significati da codificare.

Formalmente, secondo il teorema di codifica di Shannon, la lunghezza ottimale di codifica per un simbolo x con probabilità P(x) è $L(x) = -\log_2 P(x)$ bits. Questo implica che simboli frequenti (alta probabilità) dovrebbero ricevere codifiche corte, mentre simboli rari dovrebbero avere codifiche più lunghe. Aggregando su tutti i simboli, la lunghezza media di codifica in un codice ottimale è esattamente uguale all'entropia: $\langle L \rangle = \Sigma_x P(x) L(x) = H(X)$.

Evidenze empiriche suggeriscono che le lingue naturali approssimano questo principio di efficienza. Parole più frequenti tendono a essere più corte: nella maggior parte delle lingue esiste una correlazione negativa robusta tra frequenza lessicale e lunghezza in fonemi o lettere. Quantitativamente, questa relazione è spesso ben descritta dalla legge di Zipf-Mandelbrot: la lunghezza media delle parole con rank di frequenza r è approssimativamente $L(r) \approx a + b \log(r)$, dove a e b sono costanti dipendenti dalla lingua.

Tuttavia, vincoli addizionali oltre l'efficienza informazionale modellano la struttura lessicale. In particolare, esiste un trade-off tra compressione (parole corte per alta efficienza) e robustezza al rumore (parole più lunghe e distintive per ridurre confusabilità). Questo trade-off può essere formalizzato attraverso la teoria rate-distortion, che caratterizza la relazione ottimale tra rate di codifica e fedeltà di decodifica.

Data una sorgente di informazione X e una ricostruzione approssimata $\hat{X}$, definiamo una funzione di distorsione $d(x,\hat{x})$ che quantifica l'errore quando il simbolo x viene decodificato come $\hat{x}$. La distorsione media associata a un codice è $D = \langle d(X,\hat{X}) \rangle = \Sigma_x \Sigma_{\hat{x}} P(x,\hat{x}) d(x,\hat{x})$. La funzione rate-distortion R(D) specifica il rate minimo di codifica (in bits per simbolo) necessario per garantire che la distorsione media non ecceda D: $R(D) = \min_{P(\hat{x}|x): \langle d \rangle \leq D} I(X;\hat{X})$.

Per una data tolleranza di distorsione D, esiste un trade-off: rate più alti (codifiche più lunghe) permettono distorsione più bassa (maggiore fedeltà), mentre rate più bassi (codifiche più corte) implicano distorsione più alta. Le lingue naturali devono operare in una regione di questo spazio rate-distortion che bilancia efficienza comunicativa (rate basso) con comprensibilità robusta (distorsione bassa).

Dinamiche di Aggiornamento Bayesiano Incrementale

La comprensione linguistica procede incrementalmente: l'ascoltatore non aspetta la fine della frase per iniziare l'interpretazione, ma aggiorna continuamente le sue credenze man mano che ciascuna nuova parola arriva. Questo processo può essere modellato come aggiornamento bayesiano sequenziale.

Consideriamo un ascoltatore che mantiene una distribuzione di probabilità sulle possibili interpretazioni, rappresentata come un vettore di probabilità $p(t) = (p_1(t), p_2(t), ..., p_\square(t))$ al tempo t, dove $p_i(t)$ è la probabilità assegnata all'interpretazione i-esima. Inizialmente, prima di ricevere qualsiasi input linguistico, questa distribuzione è il prior: $p(0) = p\_prior$.

Quando la prima parola w_1 arriva al tempo t_1, l'ascoltatore aggiorna la distribuzione applicando il teorema di Bayes: $p_i(t_1) = P(w_1|H_i)\, p_i(0) / Z_1$, dove $P(w_1|H_i)$ è la verosimiglianza della parola w_1 sotto l'interpretazione H_i, e $Z_1 = \Sigma_\square P(w_1|H_\square)\, p_\square(0)$ è la costante di normalizzazione.

Quando la seconda parola w_2 arriva al tempo t_2, l'aggiornamento procede similmente ma usando la distribuzione aggiornata dopo la prima parola come nuovo prior: $p_i(t_2) = P(w_2|w_1,H_i)\ p_i(t_1) / Z_2$. Assumendo indipendenza condizionale delle parole data l'interpretazione (un'assunzione semplificatrice), abbiamo $P(w_2|w_1,H_i) = P(w_2|H_i)$, e l'equazione si semplifica.

In generale, dopo osservare la sequenza di parole $w_1, w_2, ..., w_n$, la distribuzione di probabilità è: $p_i(t) = (\prod_{n=1}^{t} P(w_n|H_i))\ p_i(0) / Z_n$, dove $Z_n = \Sigma_n (\prod_{n=1}^{t} P(w_n|H_n))\ p_n(0)$ assicura normalizzazione.

Questo aggiornamento incrementale implementa un processo di filtraggio bayesiano che restringe progressivamente lo spazio delle interpretazioni plausibili man mano che più evidenza si accumula. Interpretazioni incompatibili con le parole osservate ricevono probabilità decrescente e vengono progressivamente eliminate dalla considerazione attiva.

Modelli di Sorpresa e Predizione

Un aspetto cruciale della dinamica incrementale è che ciascuna nuova parola può essere più o meno sorprendente dato il contesto precedente. La sorpresa è quantificata informazionalmente come il negativo del logaritmo della probabilità predittiva: $\text{Sorpresa}(w_n) = -\log_2 P(w_n|w_1,...,w_{n-1})$.

Parole altamente predicibili dato il contesto (alta probabilità predittiva) generano bassa sorpresa, mentre parole impredette (bassa probabilità predittiva) generano alta sorpresa. La sorpresa è misurata in bits e può essere interpretata come la quantità di informazione nuova fornita dalla parola.

La probabilità predittiva $P(w_n|w_1,...,w_{n-1})$ può essere computata marginalizzando su tutte le possibili interpretazioni: $P(w_n|w_1,...,w_{n-1}) = \Sigma_i\ P(w_n|H_i)\ p_i(t-1)$, dove $p_i(t-1)$ è la probabilità dell'interpretazione H_i dopo aver osservato le prime t-1 parole.

Evidenze psicolinguistiche robuste dimostrano che la sorpresa predice precisamente il costo di elaborazione delle parole: parole con alta sorpresa mostrano tempi di lettura aumentati, fissazioni oculari prolungate, e componenti ERP amplificate (particolarmente N400 per sorpresa semantica). Questo collegamento diretto tra una quantità teorica informazionale e misure comportamentali e neurali fornisce validazione empirica forte per modelli basati su predizione bayesiana.

La sorpresa cumulativa attraverso un enunciato, chiamata anche surprisal totale o teoria dell'informazione della complessità di processing, predice la difficoltà complessiva dell'enunciato: $\text{Difficoltà} = \Sigma_n\ \text{Sorpresa}(w_n) = -\Sigma_n\ \log_2 P(w_n|w_1,...,w_{n-1}) = -\log_2 P(w_1,...,w_n)$.

Questa quantità è esattamente il negativo del logaritmo della probabilità dell'intero enunciato, fornendo una misura unificata di complessità che integra contributi lessicali, sintattici e semantici.

Parte III: Binding Temporale, Capacità della Memoria e Integrazione Computazionale

Il Problema del Binding e la Soluzione attraverso Sincronizzazione Temporale

Il problema del binding rappresenta una delle sfide computazionali più fondamentali nelle neuroscienze cognitive: come può il sistema nervoso rappresentare simultaneamente molteplici oggetti o strutture complesse mantenendoli distinti, senza che le loro features si confondano? Nel contesto della comprensione linguistica, questo problema emerge con particolare acutezza nella necessità di mantenere distinte molteplici strutture argomentali o proposizioni quando si processano frasi complesse con embeddings multipli.

Consideriamo una frase che descrive due eventi simultanei con partecipanti distinti: "Il gatto insegue il topo mentre il cane osserva il bambino". Per comprendere correttamente questa frase, il sistema cognitivo deve mantenere due strutture argomentali separate: Struttura_1 = (INSEGUIRE, agente: GATTO, paziente: TOPO) e Struttura_2 = (OSSERVARE, agente: CANE, paziente: BAMBINO). Il problema del binding è: come può il cervello rappresentare questi binding senza confondere gli agenti e i pazienti attraverso le due strutture, evitando confusioni erronee come (INSEGUIRE, agente: CANE, paziente: TOPO)?

La soluzione proposta attraverso sincronizzazione temporale è che elementi che devono essere legati insieme scaricano sincronicamente in una fase comune delle oscillazioni gamma, mentre elementi che devono rimanere distinti scaricano in fasi diverse. Formalizziamo questo meccanismo considerando popolazioni neuronali che rappresentano ciascun concetto lessicale e ciascun ruolo tematico.

Sia N_GATTO la popolazione neuronale che rappresenta il concetto GATTO, con attività oscillatoria caratterizzata da fase φ_GATTO(t). Similmente, definiamo N_TOPO con fase φ_TOPO(t), N_CANE con fase φ_CANE(t), N_BAMBINO con fase φ_BAMBINO(t), N_AGENTE con fase φ_AGENTE(t), e N_PAZIENTE con fase φ_PAZIENTE(t). Per rappresentare il binding della Struttura_1, le fasi delle popolazioni componenti devono soddisfare relazioni di sincronizzazione: φ_GATTO(t) ≈ φ_AGENTE(t) ≈ φ_1 e φ_TOPO(t) ≈ φ_PAZIENTE(t) ≈ φ_1, dove φ_1 è una fase comune assegnata alla Struttura_1.

Simultaneamente, per rappresentare Struttura_2 in modo distinto, le sue componenti devono sincronizzarsi a una fase diversa: φ_CANE(t) ≈ φ_2 e φ_BAMBINO(t) ≈ φ_2, dove φ_2 è sufficientemente separata da φ_1 per prevenire confusione: $|\varphi_1 - \varphi_2| > \Delta\varphi\_min$, dove Δφ_min è la separazione di fase minima richiesta per discriminabilità, tipicamente dell'ordine di quarantacinque-novanta gradi.

Questo schema di binding può essere implementato attraverso architetture neurali con connessioni specifiche. Consideriamo un modello semplificato dove le popolazioni concettuali (GATTO, TOPO, etc.) proiettano a un layer intermedio di "pools di binding" che implementano le strutture argomentali. Ciascun pool di binding è una sub-rete di neuroni eccitatori reciprocamente connessi che, quando attivati, oscillano coerentemente in fase gamma.

Le equazioni che governano l'attività del pool di binding k sono: τ dr□/dt = -r□ + F(Σ□ w□□ r□ + I_{ex}□,□ - I_i□□$_i$b,□), dove r□ è il tasso di firing medio del pool, τ è la costante di tempo, F è una funzione di trasferimento non-lineare, w□□ sono pesi sinaptici da popolazioni

concettuali al pool, e $I_{i□□ib,□} = Σ□≠□ g□□ r□$ rappresenta inibizione dai pool competitivi, implementando un meccanismo winner-take-all che assicura che pools diversi operino in fasi distinte.

La dinamica oscillatoria emerge quando incorporiamo un delay nel feedback inibitorio. Modifichiamo l'equazione includendo un termine di inibizione ricorrente ritardata: $τ\, dr□/dt = -r□ + F(I_{ex□,□} - w\_inh \cdot r□(t - τ\_delay))$, dove w_inh è la forza dell'auto-inibizione e $τ\_delay$ è il ritardo sinaptico. Questa equazione delay-differenziale può esibire soluzioni oscillatorie con frequenza determinata approssimativamente da $f ≈ 1/(2τ\_delay)$. Per $τ\_delay ≈ 12$ millisecondi, otteniamo $f ≈ 42$ hertz, nel mezzo della banda gamma.

Limiti di Capacità e Vincoli sulla Complessità Rappresentazionale

Un aspetto cruciale del binding attraverso separazione di fase è che impone limiti quantitativi sul numero di strutture che possono essere mantenute simultaneamente. Se rappresentiamo N strutture distinte attraverso N fasi gamma separate, la risoluzione di fase limita N.

Lo spazio delle fasi gamma è circolare con periodo $2π$ radianti. Se richiediamo separazione minima $Δφ\_min$ tra fasi adiacenti per discriminabilità, il numero massimo di strutture distinguibili è $N\_max = ⌊2π / Δφ\_min⌋$, dove $⌊·⌋$ denota la funzione floor. Per $Δφ\_min = π/4$ radianti (quarantacinque gradi), otteniamo $N\_max = ⌊8⌋ = 8$ strutture.

Questo limite teorico corrisponde notevolmente al limite classico della capacità della memoria di lavoro di "sette più o meno due" items identificato da George Miller. La convergenza non è coincidenza: suggerisce che il limite di Miller potrebbe riflettere vincoli fondamentali dell'architettura neurale di binding attraverso sincronizzazione temporale.

Possiamo raffinare questa analisi considerando che la discriminabilità tra fasi dipende dal rapporto segnale-rumore. In presenza di noise neuronale, due strutture rappresentate con fasi $φ_1$ e $φ_2$ possono essere confuse con probabilità che dipende dalla loro separazione di fase e dall'ampiezza del noise. Modellando il noise come fluttuazioni gaussiane della fase con deviazione standard $σ\_φ$, la probabilità di confusione tra due strutture è approssimativamente: $P\_confusione ≈ \exp(-(φ_1 - φ_2)^2 / (2σ\_φ^2))$.

Per mantenere probabilità di confusione sufficientemente bassa, per esempio $P\_confusione < 0{,}01$, richiediamo separazione di fase $|φ_1 - φ_2| > 2{,}58\, σ\_φ$ (corrispondente a 99% di intervallo di confidenza). Se il noise di fase è $σ\_φ ≈ π/12$ radianti (quindici gradi), allora la separazione minima richiesta è $Δφ\_min ≈ 2{,}58 \cdot π/12 ≈ 0{,}67$ radianti ≈ 38 gradi, e la capacità massima diventa $N\_max ≈ ⌊2π / 0{,}67⌋ ≈ 9$ items, consistente con il limite superiore della stima di Miller.

Gerarchie Temporali e Accoppiamento Multi-Scala

La comprensione linguistica opera naturalmente a molteplici scale temporali: fonemi e morfemi nell'ordine di decine di millisecondi, parole nell'ordine di centinaia di millisecondi, sintagmi e clausole nell'ordine di secondi. Questa struttura gerarchica temporale può essere implementata attraverso accoppiamenti cross-frequenza tra oscillazioni a diverse bande.

Consideriamo una gerarchia di tre livelli: oscillazioni gamma (30-100 Hz) per features e morfemi, oscillazioni theta (4-8 Hz) per parole e sintagmi, oscillazioni delta (1-4 Hz) per clausole e frasi. La relazione gerarchica è implementata attraverso modulazione di ampiezza cross-frequenza: la fase delle oscillazioni lente modula l'ampiezza delle oscillazioni rapide.

Formalizziamo questo attraverso un modello di oscillatori accoppiati gerarchicamente. La fase dell'oscillazione delta evolve secondo: $d\varphi_\delta/dt = \omega_\delta + \varepsilon_\delta(t)$, dove $\omega_\delta \approx 2\pi \cdot 2$ Hz = 4π rad/s è la frequenza angolare delta e $\varepsilon_\delta(t)$ rappresenta perturbazioni stocastiche.

La fase dell'oscillazione theta è accoppiata alla fase delta attraverso: $d\varphi_\theta/dt = \omega_\theta + \alpha_{\delta\theta} \cos(\varphi_\delta - \varphi_\theta) + \varepsilon_\theta(t)$, dove $\omega_\theta \approx 2\pi \cdot 6$ Hz è la frequenza theta, e $\alpha_{\delta\theta}$ è la forza di accoppiamento delta-theta. Questo termine di accoppiamento sincronizza theta a delta, ma non con rapporto di frequenze integer: piuttosto, cicli theta multipli si annidano all'interno di ciascun ciclo delta.

Similmente, gamma è accoppiato a theta: $d\varphi_\gamma/dt = \omega_\gamma + \alpha_{\theta\gamma} \cos(\varphi_\theta - \varphi_\gamma) + \varepsilon_\gamma(t)$, con $\omega_\gamma \approx 2\pi \cdot 40$ Hz.

L'accoppiamento di ampiezza è implementato modulando l'ampiezza delle oscillazioni rapide in funzione della fase delle oscillazioni lente: $A_\theta(t) = A_{\theta,0} (1 + m_{\delta\theta} \cos(\varphi_\delta(t) - \psi_{\delta\theta}))$, $A_\gamma(t) = A_{\gamma,0} (1 + m_{\theta\gamma} \cos(\varphi_\theta(t) - \psi_{\theta\gamma}))$, dove $A_{\theta,0}$ e $A_{\gamma,0}$ sono ampiezze baseline, $m_{\delta\theta}$ e $m_{\theta\gamma}$ sono indici di modulazione ($0 \leq m \leq 1$), e $\psi_{\delta\theta}$, $\psi_{\theta\gamma}$ sono fasi preferite della modulazione.

Questo sistema di equazioni accoppiate produce una cascata di modulazioni temporali: oscillazioni delta organizzano finestre temporali nell'ordine del secondo, corrispondenti a frasi; all'interno di ciascuna finestra, oscillazioni theta organizzano sub-finestre nell'ordine di centinaia di millisecondi, corrispondenti a parole; all'interno di ciascuna sub-finestra, oscillazioni gamma organizzano micro-finestre nell'ordine di decine di millisecondi, corrispondenti a features e morfemi.

Modelli di Attractor Dynamics per Rappresentazioni Semantiche

Un framework complementare per comprendere come il cervello mantiene e manipola rappresentazioni semantiche è fornito dalla teoria degli attractor networks. Un attractor è uno stato stabile di una rete neuronale dinamica verso il quale la rete converge da una varietà di condizioni iniziali.

Consideriamo una rete neuronale ricorrente completamente connessa con N neuroni, dove lo stato del neurone i è rappresentato dalla sua attività $x_i(t)$. L'evoluzione temporale è governata da: $\tau\, dx_i/dt = -x_i + \Sigma_j w_{ij} f(x_j) + I_i$, dove τ è la costante di tempo neuronale, w_{ij} è il peso sinaptico dalla neurone j al neurone i, f è una funzione di trasferimento non-lineare (tipicamente sigmoide o tangente iperbolica), e I_i è input esterno.

Gli stati di equilibrio della rete sono soluzioni dell'equazione: $x_i^* = \Sigma_j w_{ij} f(x_j^*) + I_i$. Questi stati possono essere classificati in base alla loro stabilità. Uno stato x^* è un attractor stabile se piccole perturbazioni decadono, riportando la rete a x^*. Matematicamente, questo richiede che tutti gli autovalori della matrice jacobiana $J_{ij} = \partial f_i/\partial x_j|_{\{x^*\}}$ abbiano parte reale negativa.

Per memorizzare pattern semantici come attractors, possiamo usare una regola di apprendimento hebbiana. Supponiamo di voler memorizzare M pattern $\{\xi^\mu\}_{\mu=1}^M$, dove $\xi^\mu = (\xi_1^\mu, \xi_2^\mu, ..., \xi_N^\mu)$ è un vettore binario che rappresenta il pattern μ. I pesi sinaptici sono impostati secondo la regola di Hopfield: $w_{ij} = (1/N) \sum_\mu \xi_i^\mu \xi_j^\mu$ per $i \neq j$, e $w_{ii} = 0$.

Con questa scelta di pesi, i pattern memorizzati $\{\xi^\mu\}$ sono (approssimativamente) stati di equilibrio stabili della dinamica della rete. Quando la rete viene inizializzata in uno stato vicino a uno dei pattern memorizzati, la dinamica converge verso quel pattern, implementando un processo di completamento di pattern o retrieval associativo.

La capacità di memoria della rete di Hopfield è limitata. Per N neuroni, il numero massimo di pattern che possono essere memorizzati affidabilmente è approssimativamente $M_{max} \approx 0{,}14\ N$. Oltre questo limite, emerge interferenza catastrofica: i pattern memorizzati non sono più stabili e la rete converge verso stati spurii che sono miscele dei pattern memorizzati.

Possiamo derivare questo limite analizzando la stabilità dei pattern memorizzati. Consideriamo il pattern ξ^1 e valutiamo il campo locale ricevuto dal neurone i quando la rete è nello stato ξ^1: $h_i = \sum_j w_{ij} \xi_j^1 = (1/N) \sum_j \sum_\mu \xi_i^\mu \xi_j^\mu \xi_j^1 = \xi_i^1 + (1/N) \sum_j \sum_{\mu \neq 1} \xi_i^\mu \xi_j^\mu \xi_j^1$.

Il primo termine è il segnale, proporzionale al pattern target. Il secondo termine è noise dovuto a interferenza dagli altri pattern memorizzati. Per pattern casuali non correlati, questo termine di noise è approssimativamente gaussiano con varianza $\sigma^2 \approx M/N$. Il rapporto segnale-rumore è quindi $SNR \approx 1/\sigma \approx \sqrt{(N/M)}$.

Per retrieval affidabile, richiediamo $SNR > SNR_{min} \approx 2\text{-}3$. Questo implica $M < N/SNR_{min}^2 \approx N/9 \approx 0{,}11\ N$, consistente con la stima empirica di 0,14 N.

Integrazione Multi-Livello: Dal Simbolico al Neurale

Abbiamo ora sviluppato formalismi matematici dettagliati a molteplici livelli esplicativi: inferenza bayesiana al livello computazionale, dinamiche oscillatorie al livello algoritmico, e attractor dynamics al livello implementazionale neurale. La sfida finale è specificare come questi livelli si connettono.

Un approccio è concepire le distribuzioni di probabilità bayesiane come implementate attraverso pattern di attivazione neuronale in reti attractor. Specificamente, la probabilità posteriore P(H|D) può essere rappresentata attraverso il tasso di firing medio di una popolazione neuronale dedicata all'ipotesi H.

Formalmente, sia r_H il tasso di firing della popolazione che rappresenta l'ipotesi H. Postuliamo una codifica probabilistica: $r_H = r_{max} \cdot P(H|D)$, dove r_{max} è il tasso di firing massimo. L'aggiornamento bayesiano $P(H|D) \propto P(D|H)\ P(H)$ può quindi essere implementato attraverso dinamiche neuronali che computano prodotti e normalizzazioni.

Consideriamo una rete con popolazioni $\{N_H\}$ rappresentanti ipotesi e popolazioni $\{N_D\}$ rappresentanti evidenze. Le connessioni sinaptiche w_{HD} dalla popolazione evidenza D alla popolazione ipotesi H codificano la verosimiglianza: $w_{HD} \propto P(D|H)$. Quando l'evidenza D

è attivata con intensità I_D, la popolazione ipotesi H riceve input: I_H = Σ_D w_HD I_D + I_prior,H, dove I_prior,H codifica il prior P(H).

L'attività risultante della popolazione H, dopo normalizzazione attraverso inibizione divisiva globale, approssima la probabilità posteriore: r_H = f(I_H) / Σ_H' f(I_H'), dove f è una funzione di trasferimento non-lineare. Per f(x) = exp(βx) con β → ∞, questa formula converge esattamente alla regola di Bayes.

Le oscillazioni neurali entrano in questo schema come meccanismo per implementare binding temporale e multiplexing. Diverse ipotesi H possono essere rappresentate in fasi gamma diverse, permettendo rappresentazione simultanea di distribuzioni di probabilità su spazi di ipotesi multi-dimensionali. La gerarchia temporale delle oscillazioni implementa la gerarchia della struttura linguistica: oscillazioni delta organizzano la segmentazione frasale, theta organizza il chunking lessicale e sintattico, gamma implementa il binding di features in rappresentazioni lessicali e il binding di argomenti in strutture proposizionali.

Equazioni Master per Dinamiche Integrate

Possiamo sintetizzare l'architettura completa attraverso un sistema di equazioni differenziali stocastiche accoppiate che operano a scale temporali multiple.

Al livello più rapido, le dinamiche di singoli neuroni all'interno di popolazioni oscillatorie sono descritte da modelli integrate-and-fire con noise: τ_m dV_i/dt = -(V_i - V_rest) + R(I_syn,i(t) + I_osc,i(t) + √(2D) η_i(t)), dove I_osc,i(t) = A_γ(t) cos(ω_γ t + φ_i) implementa la componente oscillatoria con ampiezza modulata dal ritmo theta/delta, e η_i(t) è noise gaussiano bianco con ⟨η_i(t)η_j(t')⟩ = δ□ δ(t-t').

Al livello intermedio, l'evoluzione delle fasi oscillatorie è governata da equazioni di Kuramoto accoppiate gerarchicamente: dφ_δ/dt = ω_δ + K_δ ∫ dx g(x - φ_δ) ρ_δ(x,t) + ξ_δ(t), dφ_θ/dt = ω_θ + K_θ ∫ dx g(x - φ_θ) ρ_θ(x,t) + α_δθ cos(φ_δ - φ_θ) + ξ_θ(t), dφ_γ/dt = ω_γ + K_γ ∫ dx g(x - φ_γ) ρ_γ(x,t) + α_θγ cos(φ_θ - φ_γ) + ξ_γ(t), dove ρ_i(x,t) è la densità di fase degli oscillatori nella banda i, g(x) = sin(x) è la funzione di accoppiamento, e ξ_i(t) sono termini di noise che catturano fluttuazioni.

Al livello più lento, l'evoluzione delle distribuzioni di probabilità su interpretazioni è governata dall'equazione di Fokker-Planck: ∂P(H,t)/∂t = -∂/∂H [v(H,t) P(H,t)] + D ∂²P(H,t)/∂H², dove v(H,t) = ∂/∂H log P(D|H) è il campo di drift determinato dalla verosimiglianza dell'evidenza accumulata, e D è il coefficiente di diffusione che cattura incertezza.

Questi tre livelli sono accoppiati: le distribuzioni di probabilità P(H,t) determinano quali popolazioni neuronali sono attivate e con quale intensità; l'attività neuronale determina le ampiezze e la coerenza delle oscillazioni; le oscillazioni organizzano temporalmente il binding e il chunking che permette rappresentazioni strutturate complesse necessarie per mantenere ipotesi interpretative ricche. La risoluzione numerica di questo sistema di equazioni accoppiate multi-scala costituisce un programma di ricerca computazionale ambizioso ma fattibile con metodi moderni. Simulazioni di questo tipo promettono di fornire predizioni quantitative dettagliate circa dinamiche temporali dell'elaborazione linguistica, patterns di attivazione neurale, e limitazioni di capacità che possono essere confrontate direttamente con dati sperimentali da neuroimaging, elettrofisiologia e studi

comportamentali, fornendo test critici per la teoria integrata della comprensione linguistica che abbiamo sviluppato.